

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 15, 2005, 20:42:00 ; Search time 655.99 Seconds  
(without alignments)  
5125.708 Million cell updates/sec

Title: US-09-904-994B-3  
Perfect score: 2999  
Sequence: 1 MKKKQEVVVTGPTKGDV.....KLCTSKPTSQVPLAQRITFF 568

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=x1h  
-Q=/cg21/USPTO.spool/US09904994/runat\_11022005\_145646\_25526/app\_query.fasta\_1.1102  
-DB=N\_Geneseq\_16dec04 -QFMT=faetap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blonum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=p2n -NDBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US09904994.QCGN\_1.1\_468.rumat\_11022005\_145646\_25526 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_16dec04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2999	100.0	2883	7	ADJ58237 Urease su
2	2973	99.1	2452	7	ADJ58249 UreaseXY
3	2960	98.7	2405	7	ADJ58240 UreaseXY
4	2953	98.5	2407	7	ADJ58246 UreaseXY
5	2588	86.3	2183	7	ADJ58243 UreaseXY

6	2326	77.6	1710	12	ADQ37850	Adq37850 H. blizoz
7	2326	77.6	8407	12	ADQ37847	Adq37847 H. blizoz
8	2297	76.6	1710	2	AAQ04329	AAQ04329 Probe for
9	2297	76.6	1710	8	ACA34557	ACA34557 Prokaryot
10	2297	76.6	1815	2	AAK14230	AAK14230 H. pylori
11	2297	76.6	2735	2	AAT44351	AAT44351 H. pylori
12	2291	76.4	1710	9	AAI60579	AAI60579 Helicobac
13	2286	76.2	1717	8	ABA00817	ABA00817 H. pylori
14	2258	75.3	1719	8	ABA00816	ABA00816 H. felis
15	2250	75.0	2767	2	AAQ12485	AAQ12485 DNA encod
16	2244	74.8	2619	2	AAQ09180	AAQ09180 Helicobac
17	2244	74.8	2619	2	AAT45680	AAT45680 H. felis
18	2241	74.7	2619	2	AAQ75319	AAQ75319 Urease ur
19	2233	74.5	2385	2	AAV62460	AAV62460 Helicobac
20	1922	64.1	1704	8	ACA23584	ACA23584 Prokaryot
21	1914	63.8	1698	8	ACA45232	ACA45232 Prokaryot
22	1907	63.6	1704	8	ACA25712	ACA25712 Prokaryot
23	1897	63.3	1704	8	ACA26184	ACA26184 Prokaryot
24	1832	61.1	1701	8	ACA42689	ACA42689 Prokaryot
25	1832	61.1	3919	13	ADT05424	ADT05424 Haemophil
26	1832	61.1	85814	13	ADT05644	ADT05644 Haemophil
27	1828	61.0	1701	11	ABD11712	ABD11712 Pseudomon
28	1826	60.9	1719	8	ACA34156	ACA34156 Prokaryot
29	1826	60.9	110000	2	AAT42063_05	Continuation (6 of
30	1825	60.9	1701	8	ACA35953	ACA35953 Prokaryot
31	1821	60.7	1878	11	ACH96250	ACH96250 Klebsiell
32	1818.5	60.6	1710	10	ADF01572	ADF01572 Bacteri
33	1818.5	60.6	2341	8	ACA45080	ACA45080 Prokaryot
34	1818	60.6	1704	9	ADA30499	ADA30499 DNA encod
35	1817	60.6	2400	2	AAV19002	AAV19002 Klebsiell
36	1817	60.6	4768	3	AAZ92097	AAZ92097 Klebsiell
37	1815	60.5	1698	8	ACA20910	ACA20910 Prokaryot
38	1809	60.3	1707	8	ACA32769	ACA32769 Prokaryot
39	1809	60.3	1707	8	ACA32790	ACA32790 Prokaryot
40	1809	60.3	86248	10	ADC00087	ADC00087 Enterobac
41	1809	60.3	87563	9	ACD19044	ACD19044 E. coli 0
42	1808	60.3	8729	9	ACD19244	ACD19244 E. coli 0
43	1797	59.9	1701	8	ACA44198	ACA44198 Prokaryot
44	1789.5	59.7	1716	4	AAH52362	AAH52362 S. epider
45	1789.5	59.7	2439	8	ACA46520	ACA46520 Prokaryot

## ALIGNMENTS

RESULT 1	
ADJ58237	
ID	ADJ58237 standard; DNA; 2883 BP.
XX	
AC	ADJ58237;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Urease subunit polypeptide complex encoding sequence.
XX	
KW	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
XX	
OS	Helicobacter felis.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	206..886
FT	/*tag= a
FT	/product= "urease polypeptide complex"
FT	897..2603
FT	/*tag= b
FT	/product= "urease polypeptide complex"
XX	
PN	EP1176192-A2.
XX	
PD	30-JAN-2002.
XX	
PF	11-JUL-2001; 2001BP-00202666.
XX	
PR	17-JUL-2000; 2000BP-00202565.



QY 561 LeuAlaGlnArgIYThrPhePhe 568  
 DB 2577 CTAGCCAGCGCTACACTTCTTC 2600

RESULT 2  
 ID ADJ58249 standard; DNA; 2452 BP.  
 AC ADJ58249;  
 DT 06-MAY-2004 (first entry)  
 XX ureaseX subunit encoding sequence #4.  
 DE ureaseX subunit encoding sequence #4.  
 KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.  
 XX Helicobacter felis.  
 OS Helicobacter felis.  
 FH Key Location/Qualifiers  
 FT CDS 48..728  
 FT /\*tag= a  
 FT /product= "urease protein"  
 FT CDS 739..2445  
 FT /\*tag= b  
 FT /product= "urease protein"

EP1176192-A2.  
 PD 30-JAN-2002.  
 PF 11-JUL-2001; 2001EP-00202666.  
 XX 17-JUL-2000; 2000EP-00202665.  
 PA (ALKU ) AKZO NOBEL NV.  
 PI Kusters JG, Catolet G;  
 DR WPI; 2002-124384/17.  
 DR P-PSDB; ADJ58250, ADJ58251.  
 XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in the diagnosis of Helicobacter felis infections and in the preparation of vaccines.  
 PT Disclosure; SEQ ID NO 13; 76bp; English.  
 XX The present invention relates to a novel Helicobacter felis urease X and Y subunit polypeptides and immunogenic fragments. The polypeptides are used in the manufacture of vaccines against Helicobacter felis infections and in diagnostic tests to detect antibodies against Helicobacter felis.  
 CC Helicobacter felis is difficult to grow so it is more convenient to use the expression products of the genes encoding the urease X and Y subunits in the manufacture of vaccines. The present sequence represents an ureaseX subunit of the invention.

XX Sequence 2452 BP; 746 A; 568 C; 590 G; 548 T; 0 U; 0 Other;  
 SQ

Alignment Scores:  
 Pred. No.: 4.56e-261 Length: 2452  
 Score: 2973.00 Matches: 562  
 Percent Similarity: 99.65% Conservative: 4  
 Best Local Similarity: 98.94% Mismatches: 2  
 Query Match: 99.13% Indels: 0  
 DB: 7 Gaps: 0

US-09-904-994B-3 (1-568) x ADJ58249 (1-2452)

QY 1 MetLysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20  
 DB 739 ATGAAATGAAAAAACAAGATATGTAATACCTACGACCCACACGCGATTAAGTG 798

QY 21 ArgLeuGlyAspThrAspLeuThrAlaGluValGlnHisAspTyrThrThrGlyGlu 40  
 DB 799 CGCTTAGAGATACGATCTTTGGCAGAAAGTAAACATGACTATACCATCTAGGCGAA 858

QY 41 GLeuLysPheGlyAlaGlyLysThrLLeaGluGlyMetGlyGlnSerAsnSerPro 60  
 DB 859 GAACCTAAATTCGGTGCAGGTAAACCTATCCGTAGGGTATGGGTCAAGCAATAGCCCA 918

QY 61 AspGluAsnThrLeuAspLeuValLLeaThrAsnAlaMetLLeaAspTyrThrGlyLe 80  
 DB 919 GATGAAAAACCTTAGATTTTGTATGATCACCACCGCATGATTAATGATCACCAGGATTT 978

QY 81 TyrLysAlaAspLLeaGlyLysAsnGlyLysLLeaGlyLLeaGlyLysAlaGlyAsn 100  
 DB 979 TACAAGCCGACATTGGCATTAATAAATGCAAAATGCAATGCGATTGGCAGGGAAC 1038

QY 101 LysAspMetGlnAspGlyValSerProHisMetValAlaGlyAlaGlyThrGluAlaLeu 120  
 DB 1039 AAGGACATGCAAGATGGCGTAAAGCCCTCATATGGTGGTGGTGGCAGAAAGCACTA 1098

QY 121 AlaGlyGlyMetLLeaThrAlaGlyGlyLLeaAspSerHisThrHisPheLeuSer 140  
 DB 1099 GCAGGGAGAGATGATTAATTAACCGTGGGGGATGATTCACACCCACTTCTCTCT 1158

QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThr 160  
 DB 1159 CCACAACAATTCCTTACCGCTTAGCCCAATGGCCTTCAACAAATGTTGGGCGGCACA 1218

QY 161 GlyProValAspGlyThrAsnAlaThrThrLLeaThrProGlyLysThrPAsnLLeaHisArg 180  
 DB 1219 GGCCCGTAGATGCGAGAAATGCGATACCATCTCCGCGCAATGGAACTTCACCCGC 1278

QY 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200  
 DB 1279 ATGTTGGCGCGCAGAAAGATATTCTATGATGTGGCTTTTGGCCAAAGGCAATGAC 1338

QY 201 SerSerLysLysGlnLeuValGlnGlnValGlnAlaGlyAlaGlyPheLysLeuHis 220  
 DB 1339 TCTAGTAAACCACTGTGTAACAAGTAGAACGGGCGGATGTTTAAATGTGAT 1398

QY 221 GluAspTrpGlyThrThrProSerAlaLLeaAspHisCysLeuSerValAlaAspGluTyr 240  
 DB 1399 GAACATCGGGCACAATCTCAAGTCGATCGATACCTGTTGACCGTAGCAGATGATAC 1458

QY 241 AspValGlnValCysLLeaThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260  
 DB 1459 GATGTGCAAGTTGTATACACACCGATACGCTCAATAGGACGATGTATGATGACACC 1518

QY 261 LeuAsnAlaMetAsnGlyArgAlaLLeaAlaThrHisLLeaGluGlyAlaGlyGly 280  
 DB 1519 CTAAATGCAATGAACGGGCGGCGCATCGCTTACCAACATGAGGAGCGGGTGAAGA 1578

QY 281 HisSerProAspValLLeaThrMetAlaGlyLysLeuAsnLLeaProSerSerThrThr 300  
 DB 1579 CACACCTGATGTATACCATGCGATGCGAGGAGTGAATTTCACTCTCTCCACAAACC 1638

QY 301 ProThrLLeaProTyrThrLLeaThrValAlaGluHisLeuAspMetLeuMetThrCys 320  
 DB 1639 CCTACTATCCCTTATACCATTAATATACGTTGCAAGAACCTTAGACATGCTTATACCTGC 1698

QY 321 HisHisLeuAspLysArgLLeaGluLysLeuGlnPheSerGlnSerArgLLeaArgPro 340  
 DB 1699 CACCACTGATGAACCGATCGGAGGATCTCCAAATTTCTCAAGCCGATCGGCCCC 1758

QY 341 GlySerLLeaAlaGluAspValLeuHisAspMetGlyValLLeaAlaMetThrSerSer 360  
 DB 1759 GGCTCTATCGCGCTGAAGATGTCTCCATGATATCGGTGATCGGAGTACAAAGTTCC 1818

QY 361 AspSerGlnAlaMetGlyArgAlaGlyGluValLLeaProAspThrTrpGlnThrAlaAsp 380  
 DB 1819 GATTTCGAAGCAAAAGGGCGGCTGGGGAAGTGATTTCTTGAACCTTGGCAAACTGCAGAC 1878

QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400

Db	1879	AAAGATATAAAGAAATTGGTAACCTTCCTGAAGATGCTGACGATTAATGACAATTCCGC	1938
QY	401	ILelysArgTyrlIeserIysTyThrIleasnProalaleuthRhIegIValserGlu	420
Db	1939	ATCAACAGCCTTAATCTCCAAATATACACCTTAATCCCGCTTGACCACATGGCGTACGACG	1998
QY	421	TyrlIegIserValGIguIuGIuGIlyValIleIleIleAspleuValIValITPasnProalaph	440
Db	1999	TATATCGCGCTGTGGAGAGAGGCGAGATCCGCCACTTGATGGTGATGGAAATCTGCTTTT	2058
QY	441	PheGIyValIysProIysIleValIleIysGIyGImetValIlePheSerGIumetGIy	460
Db	2059	TTTGGCGTAAACCCAAATAGTGATCAAAAGCGGTATGGTGATGGTGCTCTGAATGGCG	2118
QY	461	AsperSerAsnIasSerValProThrProGlnProValTyTyrraGIuIumetPhaGIyHis	480
Db	2119	GATTCATATGGCTGTGGCCACCTTCACGCGGTATTAATTCACGGGAATATTTGGGCAT	2178
QY	481	HisGIyIysAlaIysPheAspThrSerIleThrPheValSerIysValAlaTyGIuasn	500
Db	2179	CACGGCAAGCGGAATTTGACACACGACTACCTTTTGTTCCAAAGTGCCTAAGAATAAT	2238
QY	501	GIyValIysGIuIysLeuGIyLeuGIuIuArgGlnValIeuProValIysAsnCyethaRsn	520
Db	2239	GGTGTGAAGAAAAACTAGTGTTCAGCGGCAAGGCGCTCCCGGAAATAATCGCGCTAAC	2298
QY	521	IleThrIysIysAspPheIysPheAsnAspIysThrAlaIysIleThrValIAspProIys	540
Db	2299	ATCACCAGAAAGACCTTCAGATTCAACGACACAAACTGCAGAAAATCACCGTGATCCGAA	2358
QY	541	ThrPheGIuValPheValAspGIyIysLeuCyThrSerIysProThrSerGIuValPro	560
Db	2359	ACCTTCGAGGCTTTGTGTGATGCGCAAACTGTGACCTCTAAACCACTCTGAAGTGCT	2418
QY	561	IeuAlaGlnArgTyrrThrPhePhe	568
Db	2419	TTAGCCCAACGCTACACTTTCCTTC	2442

RESULT 3	
ADJ58240	
ID	ADJ58240 standard; DNA; 2405 BP.
XX	
AC	ADJ58240;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	ureaseXY subunit encoding sequence #1.
XX	
KW	immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds
XX	
OS	Helicobacter felis.
XX	
FT	Key
FT	Location/Qualifiers
FT	1..681
FT	/*tag= a
FT	/product= "urease protein"
FT	692..2398
FT	/*tag= b
FT	/product= "urease protein"
XX	
PN	EP1176192-A2.
XX	
PD	30-JAN-2002.
XX	
PF	11-JUL-2001; 2001EP-00202666.
XX	
PR	17-JUL-2000; 2000EP-00202565.
XX	
PA	(ALKU ) AKZO NOBEL NV.
XX	
PI	Kusters JG, Cattoli G;
XX	

DR	MPI; 2002-124384/17.
XX	P-PsDB; ADJ58241, ADJ58242.
PT	Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
PT	the diagnosis of Helicobacter felis infections and in the preparation of
PT	vaccines.
XX	
PS	Disclosure; SEQ ID NO 4; 76pp; English.
CC	
CC	The present invention relates to a novel Helicobacter felis urease X and
CC	Y subunit polypeptides and immunogenic fragments. The polypeptides are
CC	used in the manufacture of vaccines against Helicobacter felis infections
CC	and in diagnostic tests to detect antibodies against Helicobacter felis.
CC	Helicobacter felis is difficult to grow so it is more convenient to use
CC	the expression products of the genes encoding the urease X and Y subunits
CC	in the manufacture of vaccines. The present sequence represents an
CC	ureaseX subunit of the invention.
XX	
SEQ	Sequence 2405 BP; 715 A; 571 C; 588 G; 531 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	6.82e-260 Length: 2405
Score:	2960.00 Matches: 560
Percent Similarity:	99.30% Conservative: 4
Best Local Similarity:	98.59% Mismatches: 4
Query Match:	98.70% Indels: 0
DB:	Gaps: 7
US-09-904-994B-3 (1-568) x ADJ58240 (1-2405)	
OY	1 Metlysmetelyslglnlutyvalasnthryrtyglprothrlysglyaspdyval 20
Db	692 ATGAAATATGAAAAACAAGATATGTAAACCTTACGGACCACCAAGCGGAATAAGTG 751
OY	21 Argfenglyasptrrraspleutrpalagluvalgluhisaptyrthrthrtfgyglu 40
Db	752 CGCTTAGAGAGTACCAGTCTTTGGCGAGAAGTAAACATGACTATACCATTAAGGGGAA 811
OY	41 Gluleulyspheglyalaglylvethrllearglglymetgylginserasnserpro 60
Db	812 GAGCTCAAATTGGCGGGGTAAAACATTCGTAGGGTATGGCTCAGACGAATAGGCCA 871
OY	61 Aspglyasnthrlleauplevalllsthraasnlametiellaeaptrythrthgilyle 80
Db	872 GATGAAAACACTTAGATTATGTGATCACCAACGCAGTATTATGACTACACCGGATTT 931
OY	81 Tyrllysalaaplilegiyilleylasngilylailehihsigilylegilylvaaalyasn 100
Db	932 TATTAAGCGGCATTTGGTATTAATAAATGGCAAAATTCATGGTATTTGGCCAAGCGGGGAAC 991
OY	101 Lysaspmetglnasrglyvaserprohiismetvalyvalgilylthyrglualeu 120
Db	992 AAAACACATGCAAGATGGCGTAAAGCCCTCATATGGTGTGGGTGGGACGAAGAACAATA 1051
OY	121 Alaigygluglymetillelsthralagilylvyleaspsesthiasthrihsphleuser 140
Db	1052 GCAGGGGAAGGTATGATTATTAACCGCTGGGGGATGATTCGCACACCACTTCTCTCT 1111
OY	141 Proglinglnpheprothrllaleualaasnilyalthrthmetphegilyglilylthr 160
Db	1112 CCCCAACAATTCCTTACCGCTTACGCCAATGGTGTTCACCAACCAATTTGGAGGGGACA 1171
OY	161 Glyprovalaspglythrnsalnlathtthrtil ethxprrogiylyvstrpaslenhi earq 180
Db	1172 GGTCGGTAGTGGCAGATGCAACATCCATCACTCCGGGCAATGGAACTTGACCCG 1231
OY	181 Metleualrgalaalagluglutyrssetmetasnaigylypheulgilylvgslyasnser 200
Db	1232 ATGTGTCGGCGAGCTGAAGAGTATTCATGAAATGTGGGCTTTTGGCCAAAGCAATATGAC 1291
OY	201 Serterlyslvglnleualglinglnvalglnlalagilylaallegiypheplyleuhs 220
Db	1292 TTCAGTAAAAAACCACTGTGTAAGAACCAATGAAACCGGGCGGATTTGGCTTTAAATTCAT 1351



```

Db      813 GAGCTCAATTGGCGGGTAAACTATCCGTGAGGGCATGGGTGAGGCAATAGTCCA 872
Qy      61 AapGluSerThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspThrGlyIle 80
Db      873 GATGAAACACCTTACATTTAGTCATCACCAACGCATGATTAATGACATACACCGGAAAT 932
Qy      81 TyrLeuAlaAspIleGlyIleLeuAsnGlyLeuIleHisGlyIleGlyLeuValAspIle 100
Db      933 TACAAAGCCGACATTTGGCATTTAAATAATGGCAAAATCATATGCGATTTGGCAAGGCAAGAAC 992
Qy      101 LysAspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu 120
Db      999 AAGGACATCAAGATGGCGTAAGCCCTCATATGTCGTGGGTGTGGGACAGAAAGCATTA 1052
Qy      121 AlaGlyGlyGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140
Db      1053 GCGGGGGAAGTATGATTTATACCGCTGGGGGATGATTCACACACCCACTTCCTCTCT 1112
Qy      141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrPheGlyGlyGlyThr 160
Db      1113 CACAAACAATTCCTTACCGCTTACGCAATGGCGTTACAAACATGTTTGGCGGTGGACA 1172
Qy      161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLeuSerThrAsnLeuHisArg 180
Db      1173 GGTCCGTGATGGACGAAATGCACTACCATCATCCGGGCAAAATGGAACCTTGACCGGC 1232
Qy      181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLeuSer 200
Db      1233 ATGTTCCGGCGAGCTGMAAGATTTCTATGAATGTGGCTTTTGGCAAAAGGCAATAC 1292
Qy      201 SerSerLeuLeuGlnLeuValGluGlnValGluAlaGlyValIleGlyPheLeuHis 220
Db      1293 TCCAGTAAACAACTTTGTAGAACAAATAGAAAGCGGCGCATCCGGCTTTAAATTGCA 1352
Qy      221 GluAspThrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr 240
Db      1353 GAAGACTGGGGCAACATCCCAAGTCGATCGATCATCTGCTGAGCGTACAGATGAAATAC 1412
Qy      241 AapValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260
Db      1413 GATGTGCAAGTTTGTATCCACACCGATAGCGGTCAATGAAGCAGTTATGTAGTGCAC 1472
Qy      261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGly 280
Db      1473 CTGAATGCCATGAACCGGCGGCGCATCCATGCTTACCACTTGAAGGAGCGGCGAGGA 1532
Qy      281 HisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr 300
Db      1533 CACTCAGCTGATGTATACACATGGCAGGCGAGCTCAATATTCTACCTCTCCACAAAC 1592
Qy      301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
Db      1593 CCCACATATCCCTATTCATTTATACGTGTGCGAAACATTTACATGCTCATGACCTGC 1652
Qy      321 HisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340
Db      1653 CACCACCTTGAATTAACGATCCGCGAGGATTTTCAATTTTCCAAAGCCGTATCCGCCCC 1712
Qy      341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360
Db      1713 GGCTCTATCGCGCTGAAGATGTGCTCCATGATATTGGCGTATTCGCGATGACAAAGCTCG 1772
Qy      361 AapSerGlnAlaMetCysIleArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAsp 380
Db      1773 GATTTCGAAGCAATGGGCGCGCTGGCGAAGTATTCCTCGAACTTGCGAGATTCGGGAGAT 1832
Qy      381 LysAsnLysLeuGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspAsn 400
Db      1833 AAGAAATAAACAAATTTGGTAAAGCTTCTCGAAGATAGTGCAGATTAACGAACTTCCGT 1892
Qy      401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420

```

```

Db      1893 ATCAACGCTACATCTCCAAATATACATTAACCCCGCTTAAACCATGGGTAAGCGAG 1952
Qy      421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValValIleThrAsnProAlaPhe 440
Db      1953 TATATCGCTCTGTGTGAAGAGGGCAAAATCGTGATTTGGTGTGTGAAATCTGCCTTT 2012
Qy      441 PheGlyValLysPProLysIleValIleLysGlyGlyMetValValPheSerGluMetGly 460
Db      2013 TTGGTGTGAACCTTAAGATTGTGATCAAAAGCGGATATGATGTCTTCTCTGAATGGGC 2072
Qy      461 AapSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480
Db      2073 GACTCCAAAGCGCTCCGTGCTTACCTCAGCCGGTTTATACCGGAAATGTTTGGGCAT 2132
Qy      481 HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn 500
Db      2133 CACGGCAAGCGCAAAATTTGACACAGCATCTTTGTGTTCCAAAGTGCCTATGAAAT 2192
Qy      501 GlyValLysGluLysLeuGlyLeuGluLysArgGlnValLeuProValLysAsnCysArgAsn 520
Db      2193 GCGCTGAAGAAAGAAACTAGGCTTACAGCGCAAGATGCTACCGGTAAACACCGCGCAAC 2252
Qy      521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys 540
Db      2253 ATCACTAAGAAAGACTTCAATTCACAAACAAAGCGGCGCATATCATCTGATCTTAAA 2312
Qy      541 ThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValPro 560
Db      2313 ACCTTGAGAGTCTTTGTGATGAGCAAACTGTGACCTTAAACCCGCTGAAAGTGCT 2372
Qy      561 LeuAlaGlnArgTyrThrPhePhe 568
Db      2373 CTAGCCAGCGCTACACTTCTTC 2396

RESULT 5
ADJ58243
ID ADJ58243 strand; DNA; 2183 BP.
XX
AC ADJ58243;
XX
DT 06-MAY-2004 (first entry)
XX
DE UreaseX subunit encoding sequence #2.
XX
KW immunogenic fragment; antibacterial; Vaccine; Helicobacter felis; ds.
OS Helicobacter felis.
XX
FH Key Location/Qualifiers
FT CDS 3..683
FT /tag= a
FT /product= "urease protein"
FT CDS 694..2181
FT /tag= b
FT /product= "urease protein"
XX
PN EP1176192-A2.
XX
PD 30-JAN-2002.
XX
PF 11-JUL-2001; 2001EP-00202666.
XX
PR 17-JUL-2000; 2000EP-00202565.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Kusters JG, Catcoli G;
XX
XX WPI: 2002-124384/17.
XX DR P-PSDB; ADJ58244, ADJ58245.
XX
XX Novel Helicobacter felis urease X and Y subunit polypeptides, useful in
XX the diagnosis of Helicobacter felis infections and in the preparation of

```

PT vaccines.  
 XX  
 PS Disclosure; SEQ ID NO 7; 76pp; English.  
 XX  
 CC The present invention relates to a novel *Helicobacter felis* urease X and  
 CC Y subunit polypeptides and immunogenic fragments. The polypeptides are  
 CC used in the manufacture of vaccines against *Helicobacter felis* infections  
 CC and in diagnostic tests to detect antibodies against *Helicobacter felis*.  
 CC *Helicobacter felis* is difficult to grow so it is more convenient to use  
 CC the expression products of the genes encoding the urease X and Y subunits  
 CC in the manufacture of vaccines. The present sequence represents an  
 CC ureaseX subunit of the invention.  
 XX  
 SQ Sequence 2183 BP; 638 A; 505 C; 550 G; 490 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 5.26e-226 Length: 2183  
 Score: 2588.00 Matches: 489  
 Percent Similarity: 99.59% Conservative: 1  
 Best Local Similarity: 99.39% Mismatches: 2  
 Query Match: 86.30% Indels: 0  
 DB: Gaps: 0  
 US-09-904-994B-3 (1-568) x ADJ58243 (1-2183)  
 QY 1 Metlyemetylsynginglutyryvalasnthrytgytprothrysglyasplyval 20  
 DB 694 ATGAAAATGAAAAACAGAGATATGTAAACCTACGACCCACACAGCGGATTAAGTG 753  
 QY 21 Argleuglyasptthraspleutrpalaigluvalaiguhiasptrythrthrtgyglu 40  
 DB 754 CGCTTAGAGATACCAGATCTTGGCGAGAGTAAACATGCTATACCATTAATGGCGAA 813  
 QY 41 Gluleulysphegiyalaglylvethrileargluglymetgylginserasnsertpro 60  
 DB 814 GAGTCAAAATTTGGCGGGTAAACATATCCGTAGGGTATGGGTACAGAAATAGCCCA 873  
 QY 61 Aspglyasnthrieuapleuvalillethrasnalametilleaspttrythrthrtyle 80  
 DB 874 GATGAAACACCTTAGATTAATGATACCAACCGCATGATTAATGCTACACCGGAGAT 933  
 QY 81 Tyrtysalaaaspliegilylelyasnglylysiilehiegilylelyalaglyasn 100  
 DB 934 TATTAACCCAGCATTTGATTAATAAATGCGAAATCCATGATATTGGCAGGCGGAGAC 993  
 QY 101 Lysaspmetglnaspglyvalserprohismerivaiglyvalgylthrglualeu 120  
 DB 994 AAAGACATGCAAGATGCGTAAAGCCCTCATATGTCGTGGGTGGGCAAGAGCACTA 1053  
 QY 121 AlaglygluglymetillethralaglylyleaspserrhietrhleasleuSer 140  
 DB 1054 GCAAGGGAAGATTAATTAATTCGCTGGGGGAGATTCATTCGACACCCACTTCTCTCT 1113  
 QY 141 Proglnglnpheprothralaleuvalaasnalyvalthrthmetphegylgylgylthr 160  
 DB 1114 CCCCAACATTCCTTACCTGCTTACGCAATGCTTACACCACTTTTGGAGGTGGACA 1173  
 QY 161 Glyprovalaspglythrnaamlaethrthlethrprogllystppasnlleuhsarg 180  
 DB 1174 GGTCCGGTAGATGCGAAGATGCGACACCATCACTCCGGGCAATGGAACCTTCACCGC 1233  
 QY 181 Metleuarglaalagluglytyrsermetasnaiglypheleuglylvaglyasnser 200  
 DB 1234 ATGTGGCGCAGCTGAAGAGTATTTATGAAATGTAGCTTTTGGCAAGGCAATAGT 1293  
 QY 201 SerSerlysynglnleuvalaigluvalaigluaglyalalegilyphelyleuhs 220  
 DB 1294 TCTAGCAAAAACAACCTTGTAGAACAGTAAAGCGGCGGAGTGGCTTTAAATTCAT 1353  
 QY 221 GluaspttglythrthrproseralaleasphiascyseuSerValaiaaspglytyr 240  
 DB 1354 GAAGACTGGGGCACACACCAAGTGCATGCATGCTTGAACGTGGCGAGATGAATAC 1413

QY 241 Aspvalglnvalcysillethrasptthrasnengluvalaglytyrvalaaspsptthr 260  
 DB 1414 GATGTGAAGTTGTATTCACACGATACGCTCAATAGGCGAGGTATGTGGATGACACC 1473  
 QY 261 Leuasnalametasnenglyasgalailehialatyrhisilegilyalaglygyl 280  
 DB 1474 CTAAATCATGAACGGGCGGCGCATCTCAATCCATTAAGAGAGCGGCGAGGA 1533  
 QY 281 HiserproaspvalillethmetalaiglyleuasnilleuaproserSerThrthr 300  
 DB 1534 CACTCACTGATGTATACCATGACATGAGCGAGCTCAATATTTACCTCTCCACACC 1593  
 QY 301 Prothrlleprotyrthrthleasnthrvalaaleuhsleuaspmetleuethrcys 320  
 DB 1594 CCCACTATTCCTTATACCATTAATACGTTGCAACACTTAACATGCTCATACCTGC 1653  
 QY 321 HisleuaspplysasgyllearglyuasplenglnpheSerGlnserArgleargpro 340  
 DB 1654 CACCACTAGATTAAGCCCATCCGAGAGATTTACAAATTTCTCAAAACCGATCCGCC 1713  
 QY 341 GlyserllealalagluaspvalleuhsaspmetgylvalillealmetThrSerSer 360  
 DB 1714 GATCTATTGCGCGCTGAGATGTCTCATGATATTTGGCGGATCGGATGACTAGCTCC 1773  
 QY 361 Aspsertglnalametgylatglaiglygluvalilleproargthrttrpglntthralasp 380  
 DB 1774 GATTGCAACCAAGGGGCGCGCTGGGGAAGTGTCTTGAACTTGCAACTGGCAATGCGAGT 1833  
 QY 381 LysasnlysynglusphegiylsleuProgluaspglylysaaspaaspaasphaarg 400  
 DB 1834 AAGAAATAAAAGAAATTTGTTAAGCTTCTCGAAGATGTGCAAGTAAACGACACTTCGC 1893  
 QY 401 Ilelyahrgtyrilleserlytyrthrileasproalaethrhisiglyvalserglu 420  
 DB 1894 ATCAAAACGATATATCTCAATATACCATTAATCCCTTTGACCCATGCGGTAGCGAG 1953  
 QY 421 TyrtilegyservalagluglylvethrilealaspseuvalailetpaanproalaPhe 440  
 DB 1954 TATATCGCTCTGTGGAAGGCGCAAGATGCGCATGTTGGTGGTGAATTCGCTTT 2013  
 QY 441 PhegilyvallyspProlysiilevalillelyeglygylmetvalaPheSerGlymetGly 460  
 DB 2014 TTTCGCTGAACCTTAAGATTTGATTAAGAGTGCATGTTGCTTCTGAAATGGCC 2073  
 QY 461 AspsersnalametvalaProthrProglnglnprovalityrtyrarglumetphegylhis 480  
 DB 2074 GATCTTAACCGCTCGCGCCACGCTCAGCGGTTTATTAACGCGAAATGTTGGGCAC 2133  
 QY 481 HisgilyvalaalyspheasptthrsertillethPhe 492  
 DB 2134 CACGCAAGGCGAAATTTGAACACAGCATCATTTT 2169  
 RESULT 6  
 ADQ37850  
 ID ADQ37850 standard; DNA; 1710 BP.  
 XX  
 AC ADQ37850;  
 XX  
 DT 07-OCT-2004 (first entry)  
 XX  
 DE H. bizozeronii ureB DNA.  
 XX  
 KW Urease; urease gene cluster; urease structural gene;  
 KM urease accessory gene; ureB; *Helicobacter bizozeronii* infection;  
 XX antibacterial; gene; ds.  
 XX  
 OS *Helicobacter bizozeronii*.  
 FH Key Location/Qualifiers  
 FT 1..1710  
 FT /\*tag= a  
 FT /product= "H. bizozeronii ureB"  
 XX

PN US2004142343-A1.  
XX  
XX 22-JUL-2004.  
XX  
XX 12-AUG-2003; 2003US-00639273.  
XX  
XX 16-AUG-2002; 2002US-0404337P.  
XX  
XX (CHAN/) CHANG Y.  
XX (SIMP/) SIMPSON K W.  
XX (ZHUC/) ZHU J.  
XX  
XX Chang Y, Simpson KW, Zhu J;  
XX  
XX WPI: 2004-533502/51.  
XX P-PSDB; AD037851.  
XX  
XX Novel isolated nucleic acid molecule having urease gene cluster, and  
XX PT conferring on Helicobacter bizzozeroni ability to produce urease, useful  
XX as vaccine for preventing disease in mammals infected by H.bizzozeroni.  
XX  
XX Claim 7, SEQ ID NO 4; 40pp; English.  
XX  
XX The invention relates to an isolated nucleic acid molecule conferring on  
XX CC Helicobacter bizzozeroni an ability to produce urease, where the nucleic  
XX CC acid molecule is a urease gene cluster comprising at least one urease  
XX CC structural gene and at least one urease accessory gene. The nucleic acid  
XX CC molecule is chosen from urea, ureb, ureg, uref, ureg, ureh and urei. The  
XX CC invention also relates to an isolated protein encoded by the nucleic  
XX CC acid, a vaccine for preventing onset of disease in mammals infected by H.  
XX CC bizzozeroni comprising a nucleic acid and a carrier, and an isolated  
XX CC antibody or its binding portion raised against the nucleic acid. The  
XX CC nucleic acids, proteins and antibodies are useful for vaccinating mammals  
XX CC against onset of disease caused by infection of H. bizzozeroni, which  
XX CC involves administering the sequences. The sequences are useful for  
XX CC detecting H. bizzozeroni in a sample of tissue or body fluids which  
XX CC involves providing a nucleic acid as an antigen, providing an antibody,  
XX CC or providing a nucleotide sequence as a probe in a nucleic acid  
XX CC hybridisation assay, contacting the sample with the antigen or the probe,  
XX CC and detecting any reaction which indicates that H. bizzozeroni is  
XX CC present in the sample. This sequence represents H. bizzozeroni ureb DNA.  
XX  
XX  
SQ Sequence 1710 BP; 492 A; 436 C; 386 G; 396 T; 0 U; 0 Other;  
Alignment Scores:  
Score: 3,06e-202 Length: 1710  
Perc: 2326.00 Matches: 420  
Percent Similarity: 87.30% Conservative: 75  
Best Local Similarity: 74.07% Mismatches: 72  
Query Match: 77.56% Indels: 0  
DB: 12 Gaps: 0  
US-09-904-994B-3 (1-568) x AD037850 (1-1710)  
QY 2 LysMetIysLeuGlnGluTyrValAsnThrTyrGlyProThrIysGlyAspLysValArg 21  
DB 7 AAAATCTCTCGAAAGAAATATGTTCTATGTAATGAGCCACATACGGCGATTAAGGAGA 66  
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrThrTyrGlyGluGlu 41  
DB 67 TGGGGGAGATCCGACCTGATCTTAGAAGTCGAAACATGACGACCCACTTAATGCGAAGAA 126  
QY 42 LeuLysPheGlyValaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61  
DB 127 ATTAATTTGGTGGCGGTAAGAACATTCGCGATGGGAGGACAAACCAAGCCCAACC 186  
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81  
DB 187 AGCCACGAACTGATTTGTGCTCACTAAGCCCTGATGCTGATTAACACCGCATTTAT 246  
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyTyrGlyLysAlaGlyAsnLys 101  
DB 247 AAAGCGATATTGGCATTAATAATGCAAAATCATGCGATTTGGCAAGCAGGCAATAA 306

QY 102 AspMetGluAspGlyValSerProHisMetValValGlyValaGlyThrGluAlaLeuAla 121  
DB 307 GACATGCAAGATGGCGTTTGCACAAATCTTGGCGGGCCCTGCTACTGAGGCTTTGGCC 366  
QY 122 GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141  
DB 367 GCTGAAGGGCTGATTTGTAACAGCTGTGGATTGACACCATCATCACTTATTTCTCC 426  
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyIleThrGly 161  
DB 427 CAACAAATCCCAACAGCATTTGCGACGGAGACACACCATGATTTGGGGGAAACAGT 486  
QY 162 ProValaAspGlyThrAspAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181  
DB 487 CCAAGCTGATGGACATACCGACCTACCATCACTCCGGGGCGCTGAACTTAAACATG 546  
QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysAsnSerSer 201  
DB 547 CTCGGGCTCTGAAGAAATATGCATGAACCTTGGGCTATTGGTAAAGGAATGTCT 606  
QY 202 SerLysLysGlnLeuValGluGlnValaGlyValaIleGlyPheLysLeuHisGlu 221  
DB 607 TATGAACCTCCCTGCTGATCACTCGAAGCTGAGCCATTGGCTTTAAATCCACGAA 666  
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241  
DB 667 GACTGGGATGACACCTGACACCTTACCATTTGCTTAATGTGGCTACAAATACAT 726  
QY 242 ValGlnValCysIleHisThrAspThrValaGluValaGlyTyrValaAspAspThrLeu 261  
DB 727 GCGAAGTGGCTATCCACACCATACCTGAATGAAGCGGGCTGTGGAAACACTTTG 786  
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValaGlyIleHis 281  
DB 787 CAAGCATTTGCGGCGCATATCCACTTCCACTGAAGGTGCTGTGGCGGC 846  
QY 282 SerProAspValIleThrMetAlaGlyLeuAsnIleLeuProSerSerThrPro 301  
DB 847 GCTCGGATGATTAAGATGCTGCGAATTAATCACTCCAGCTTTCACCAACCC 906  
QY 302 ThrIleProTyrThrIleAsnThrValaGluHisLeuAspMetLeuMetCysHis 321  
DB 907 ACCATTTCTTACCGTGAATACAGAACCCACATGACATGTGATGTGTCAC 966  
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
DB 967 CACTTGGATTAACATCAAGAAAGATGTCAGTTTGTGATTTAGATTTGCCCCCA 1026  
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetSerSerSerAsp 361  
DB 1027 ACCATCGCGCTGAGAGACAAATCCACGATATGGGATTTTCTATACCGACTTGAC 1086  
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381  
DB 1087 TCCCAAGCGATGGCGGTAGAGGATGATCACCCGACATGGCAAAACACGGACAA 1146  
QY 382 AsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401  
DB 1147 AACCAAAAGAAATTTGGTGTGCTTAGAGGAAAGGCGAATATGACAACTTCGCATC 1206  
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValaSerGluTyr 421  
DB 1207 AAGCGCTACATTTCCAAATACCATCAACCCCGTATTTGCACAGGACATTTCTGAAT 1266  
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441  
DB 1267 GCGGCTCTGTAAGATGGGCAAAATTCGCGCATTTGGTCTTTGAGTCCCTCGTCTTT 1326  
QY 442 GlyValLysProLysIleValIleLysGlyIleMetValaIlePheSerGluMetGlyAsp 461  
DB 1327 GGCATTTAAACCAACATGATCATCAAGCGGATTCATCGACCTTTTCAAAATGGCGCAT 1386



```
QY 462 SerenA1aseValProthProGlnProValTyrTyrArgGluMetPheGlyHisHis 481
D 1387 GCCATGCTCTATCCACATCCCAACCCGATATTACCCGAAATGTTGGCCACCAT 1446
QY 482 G1YLYSA1ALysPheAspThrSerIleThrPheValSerIleValAlaTyrGluAsnGly 501
D 1447 GGTAAAGCAAAATTTGACACCAATATCACTTTGATCCCAAGTGGCTTAGACAAAGCC 1506
QY 502 VallysgLysLeuGlyLeuGlnValLeuProValIlyAsnTyrAsnGlnIle 521
D 1507 ATTTAAAGAGAGTGGGCTTCACAAAGAGTGGTTTGCAGTTAAATACTGCGCAACATC 1566
QY 522 ThrIlySA1AspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
D 1567 ACCAAAGAAAGACCTCAAAATTAACGATGTTACCGACACATCGAAGTCAATCTGAAGCC 1626
QY 542 PheGluValPheValAspGlyLysLeuGlySerThrSerIlyProThrSerGlnValProLeu 561
D 1627 TACAAAGTTAAAGTGATGCGAAAGAGGTTACTTCCAAAGACGCGGATTAATCAAGCTTA 1686
QY 562 AlaGlnArgTyrThrPhePhe 568
D 1687 GCACAACTCTACACTTGTTC 1707

RESULT 7
ID ADQ37847 standard; DNA; 8407 BP.
AC ADQ37847;
XX
DT 07-OCT-2004 (first entry)
XX
DE H. bizozeroni urease gene cluster, ureABIFGH.
XX
KM Urease; urease gene cluster; urease structural gene;
KM urease accessory gene; ureABIFGH; Helicobacter bizozeroni infection;
KM antibacterial; gene; ds.
XX
OS Helicobacter bizozeroni.
XX
PN US2004142343-A1.
XX
PD 22-JUL-2004.
XX
PF 12-AUG-2003; 2003US-00639273.
XX
PR 16-AUG-2002; 2002US-0404337P.
PA (CHAN/) CHANG Y.
PA (SIMP/) SIMPSON K W.
PA (ZHU/) ZHU J.
XX
PI Chang Y, Simpson KW, Zhu J;
XX
DR WPI: 2004-533502/51.
DR GENBANK; AF330621.
PT
XX
XX Novel isolated nucleic acid molecule having urease gene cluster, and
XX as vaccinia for preventing disease in mammals infected by H.bizozeroni.
XX
PS Claim 2; SEQ ID NO 1; 40bp; English.
XX
XX The invention relates to an isolated nucleic acid molecule conferring on
XX Helicobacter bizozeroni an ability to produce urease, where the nucleic
XX acid molecule is a urease gene cluster comprising at least one urease
XX structural gene and at least one urease accessory gene. The nucleic acid
XX molecule is chosen from ureA, ureB, ureC, ureD, ureE, ureF, ureG, ureH and ureI. The
XX invention also relates to an isolated protein encoded by the nucleic
XX acid, a vaccine for preventing onset of disease in mammals infected by H.
XX bizozeroni comprising a nucleic acid and a carrier, and an isolated
XX antibody or its binding portion raised against the nucleic acid. The
XX nucleic acid, proteins and antibodies are useful for vaccinating mammals
```

```
CC against onset of disease caused by infection of H. bizozeroni, which
CC involves administering the sequences. The sequences are useful for
CC detecting H. bizozeroni in a sample of tissue or body fluids which
CC involves providing a nucleic acid as an antigen, providing an antibody,
CC or providing a nucleotide sequence as a probe in a nucleic acid
CC hybridisation assay, contacting the sample with the antigen or the probe,
CC and detecting any reaction which indicates that H. bizozeroni is
CC present in the sample. This sequence represents the H. bizozeroni
CC urease gene cluster, ureABIFGH.
XX
SQ Sequence 8407 BP; 2264 A; 1937 C; 1969 G; 2237 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,56e-201 Length: 8407
Score: 2326.00 Matches: 420
Percent Similarity: 87.30% Conservative: 75
Best Local Similarity: 74.07% Mismatches: 72
Query Match: 77.56% Indels: 0
DB: 12 Gaps: 0
US-09-904-994B-3 (1-568) x ADQ37847 (1-8407)
QY 2 LysMetLysGlnGlnGlyTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21
D 3068 AAAATCTCGAAAGAAATATGTTTCTATGATGAGCCCACTAGCGCGATTAAGTGA 3127
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41
D 3128 TTGGGCGATACCGACATGATCTTAGAAGTCAATGATGACATGACATGATGCGAAGAA 3187
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGlnGlyMetGlyGlnSerAsnSerProAsp 61
D 3188 ATTAAGTTTGATGGCGGTAAACCATTCGCGATGGATGGACAAACCAAGCCAGC 3247
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyTyr 81
D 3248 AGCCACGAATCTCATCTTGCTGCTACTAAGCCCTGATCGGATTAACCGGATTTAT 3307
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
D 3308 AAAGCCGATATGGCATTAATAAATGCAAAATCCATGCGCATGGCAAGCAGCAATAAA 3367
QY 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValAlaGlyThrGlnAlaLeuAla 121
D 3368 GACATGGCAAGATGGCGTTTGCACAACTTTGGCTGGGCGCTGACTGAGGCTTTGGCC 3427
QY 122 GlyGlnGlyMetIleIleThrAlaGlyIleIleAspSerHisThrHisPheLeuSerPro 141
D 3428 GCTGAAGGCGGATGTTTACAGCTGGTGGGATTCACACCATCCACTTATTCTCC 3487
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyIleThrGly 161
D 3488 CAACAAATCCCAACAGCATTTGGCCAGCGGATCAACACATGATGGTGGGGAACAGGT 3547
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
D 3548 CCAAGCTATGGAGACTTAACGACATCAATCCATCCCGGGGCGCTGGAACCTTAAACCATG 3607
QY 182 LeuArgAlaIleGlnGlyTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
D 3608 CTCGGTCCCTGAAAGAAATATGCCATTAATCTGGGCTTATTTGGGTAAGGAATGTCT 3667
QY 202 SerLysLysGlnLeuValGlnGlnValAlaGlyAlaIleGlyPheLysLeuHisGly 221
D 3668 TATGAACCTCCCTCGCTGGTCACTCACTGAGCTGAGACCTGATTTAAATATCAGAA 3727
QY 222 AspThrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241
D 3728 GACTGGGGTACACACCTGAGGCAATCTTACATTTGTAATGGGTGACAAATACGAT 3787
QY 242 ValGlnValCysIleHisThrAspThrValAsnGlyAlaGlyTyrValAlaAspThrLeu 261
D 3788 GTGCAATGGCTATCCACACGATACCTTGAATGAAGCGGGCTGTGTGGAAGACATTTG 3847
```

```

QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaIleHisIleGluGlyValGluGlyHis 281
DB 3848 CAAAGCATTGCTGCTGACCTATCCACTTCCACACTGAGTGTGGCGGAC 3907
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrPro 301
DB 3908 GCTCCGGATGTCATTAAAGTGTGGCGAATTAAATCATCTCCAGCTTCAACAAACCC 3967
QY 302 ThrIleProTyThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321
DB 3968 ACCATTCCCTTTCACCCGTGATACAGAAAGCCGAAACATGATGATGATGTCGCCAC 4027
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 4028 CACTTGATTAATAACATCAAGAGAGATGTCACATGCTGCTGATTTCTAGATTCGCCCCCAA 4087
QY 342 SerIleAlaAlaGluAspValIleuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 4088 ACCATGCGCGTGGAGCAAACTCCAGATATGGGATTTCTCTATCAACAGCTCTGAC 4147
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
DB 4148 TCCCAAGCATGGCGGTAGCGAGTCATCACCCGCACTTGGCAACAGCGGACAA 4207
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
DB 4208 AACAAAAAGAAATTTGCTGCTGCTGAGGAAAAAGCGATATATACATTCGCCGATC 4267
QY 402 LysArgTyIleSerLysTyThrIleAsnProAlaLeuThrHisGlyValSerGluTy 421
DB 4268 AAGCGCTACATTTCCAAATATACCATCAACCCCGTATATGACACGCGCATTTCTGATAT 4327
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
DB 4328 GTCGGCTCTGTAGAAATGGGCAAAATTCGCGATTTGTGCTTGGAGTCCGCTCTTT 4387
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
DB 4388 GGCATTAAACCCCAACATGATCATCAAGGCGGATTCATGCGACTTTCTCAATGGGCGAT 4447
QY 462 SerAsnAlaSerValProThrProGlnProValTyTyArgGluMetPheGlyHisHis 481
DB 4448 GCCAATGCTCTATCCCACTCCCAACCCGCTATTTACCGCAAAATGTTGGCCACCAT 4507
QY 482 GlyLysAlaLysPheSerThrSerIleThrPheValSerLysValAlaTyTyGluAsnGly 501
DB 4508 GGTAAAGCCAAATTTGACACCAATATCCTTTGTATCCCAAGTGGCTTATGCAACGCGC 4567
QY 502 ValLysGluLysLeuGlyLeuGluArgGlnValLeuProValLysAsnCysArgAsnIle 521
DB 4568 ATTAAAGAAAGAGTGGGCTTGCAAGAGTGTTCCTCCAGTTAAATCCGCCAACAATC 4627
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
DB 4628 ACCAAAAAGACCTCAATTTCAACGATGTATACGCGACACATCGAAGTCAATCCTGAAC 4687
QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
DB 4688 TACAAAGTTAAAGTGGTGAAGAGGTTACTTCCAAAGCAGCGGATTAATCAACGCCA 4747
QY 562 AlaGlnArgTyThrPhePhe 568
DB 4748 GCACAACTCTACAACTTGTTC 4768

```

RESULT 8  
AAQ04329 standard; DNA; 1710 BP.

XX AAQ04329;

XX 24-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)

```

DT 21-SEP-1990 (first entry)
XX
DE Probe for sequence encoding part of protein with urease activity.
XX
XX Urease; probe; ss.
XX
XX Helicobacter pylori.
XX
XX WO9004030-A.
XX
XX 19-APR-1990.
XX
XX 06-OCT-1988; 88FR-00013135.
XX
XX 06-OCT-1988; 88FR-00013135.
XX
XX (INSP ) INST PASTEUR.
XX
XX (INRM ) INSERM INST NAT SANTE & RECH MED.
XX
XX Labigne A;
XX
XX MPI, 1990-147844/19.
XX
XX F-PSDB; AAR04580.
XX
XX
XX New nucleotide sequences encoding Campylobacter pylori-ureasents - and
XX derived vectors, transformants, protein, antibodies and probes, useful in
XX diagnosis, treatment and prevention of infections.
XX
XX Claim 7, Page 32; 47p; French.
XX
XX
XX Detection probes consisting of all/part of the sequence can be used to
XX detect C. pylori eg in stomach biopsy samples. The DNA can also be used
XX to express proteins for the prodn. of antibodies and to prepare vaccines.
XX See also AAQ04309, AAQ04311, AAQ04328-30 and AAR04581. (Updated on 25-MAR-
XX -2003 to correct PA field.) (Updated on 24-OCT-2003 to standardise OS
XX field)
XX
XX SQ Sequence 1710 BP; 537 A; 365 C; 387 G; 421 T; 0 U; 0 Other;
XX
XX
XX Alignment Scores:
XX Pred. No.: 1,366-199 Length: 1710
XX Score: 2297.00 Matches: 416
XX Percent Similarity: 86.77% Conservative: 76
XX Best Local Similarity: 73.37% Mismatches: 75
XX Query Match: 76.59% Indels: 0
XX DB: 2 Gaps: 0
XX
XX US-09-904-994B-3 (1-568) x AAQ04329 (1-1710)
XX
XX 2 LysMetLysLysGlnGluTyTyValAsnThrTyTyGlyProThrLysGlyAspLysValArg 21
XX 7 AAGATTACGACGAAAGAAATATGTTCTATGTAATGCTCTACTACAGCGGATTAATGAGA 66
XX
XX 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyThrTyTyGlyGluGlu 41
XX 67 TTGGCGCATACAGACTTGATGCTGTAAGAGAAATGATGATGATGATGATGATGATGATG 126
XX
XX 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
XX 127 CTTAATATTCGGTGGCGGTAAACCTTAAGAGAAAGCAATGCAATCTTAATGCGATTAAT 186
XX
XX 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspTyThrGlyIleTy 81
XX 187 AAAGAAAGCTTGATTTAATTAATCTCAACGCTTAAATCGTGATTAACCGGATTTAT 246
XX
XX 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
XX 247 AAAGCGATATATTGATTAAGATGCAAAATCGCTGCAATGCTGAAGCGGTAAACAA 306
XX
XX 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyTyThrGluAlaLeuAla 121
XX 307 GACATGCAAGATGGCGTTAAACAAATCTTACGTAGTGTCTGCTACTGAAAGCTTATGCC 366

```

QY 122 GYGLUGLWMECTIIEIETHRALAGLYGVIIEAEPSEHISTHRIAPHELEUSEPRO 141  
 DB 367 GGGAGAGTTTGATCGTACGCGTGGTGGTATGTACACACATCCACTTCAATTCACACC 426  
 QY 142 GINGINPHEPROTHRALALEUALAANGLYVALTHRTMCEPHEGLYGLYTHRG1Y 161  
 DB 427 CAACAAATCCCTACAGCTTTTGCAGAGCGGTTCACACCAATGATTTGGTGGAGAACCGGT 486  
 QY 162 PROVALAEPGLYTHRAENALATHRTIETHRPROGLYSTRPAENLEUIARGMET 181  
 DB 487 CCTGCTATGGCACTAATGACACTACTATCACTCCAGCGACAAATAATTTAAATGATG 546  
 QY 182 LEAHAGLALAGLUGLYRSEMERCEASNAVALGYPHELEUGLYYGLYASNSER 201  
 DB 547 CTCGAGCGGCTGAGAAATATCTATGAAATTTAGGTTCTTGGCTAAAGGTAACGCTTCT 606  
 QY 202 SERLYSELGSLNEUVALGLUGLVALGLVALGLYALIEGLYPHELYLEHISGLU 221  
 DB 607 AAGCATGCGAGCTTAGCGCATCAAAATGAAAGCCGGTGGATTTGGCTTTAAATTCACGAA 666  
 QY 222 ASPTTGPLYTHRTHPROSERALALEAPHSICYSLEUSERVALAIAEPGLUTYRASP 241  
 DB 667 GACTGGGGCACCACTCTTCTGCAATCAATCATCGTTAGATGTTGGCGGACAAATACGAT 726  
 QY 242 VALGLINVALCYSLIEHISTHRAEPTHRVALAENGLUAGLYTYRVALAEPSPTHRLEU 261  
 DB 727 GTGCAAGTCGCTATCCACACAGACACTTTCACACTGAAAGCGCTGTGTAGAGAACACTATG 786  
 QY 262 ASNALAMEAENGLYARGALALEHIALATYRHISILEGLIYALAGLYGLYVHS 281  
 DB 787 GCTGCTATGCTGAGACCCACTATGACACACTTTCACACTGAAAGCGCTGGGGCGGACAC 846  
 QY 282 SERPROASPVALIETHRMETALAGLYLEUENILEUPROSERSETHRTHPRO 301  
 DB 847 GCTCCTATATTTAATAAGCCGGTGAACACAACATTTCTCCCGCTTCCACTAACCCCC 906  
 QY 302 THRILEPROTYRTHRIEANTHRVALAAGLUHISLEAPMETLEUMETHCYEHIS 321  
 DB 907 ACCATCCCTTTCACCGTAATACAGAAAGACAGACACTGACATGCTTATGTGTGTCAC 966  
 QY 322 HISLEUAPLYSARGILEARGLUASPLEUGINPHESEGINSEARGILEARGPROGLY 341  
 DB 967 CACTGTGATTAAGCATTTAAAGAGATGTTCACTTCCTGATTCAGAGATCCGCCCTCAA 1026  
 QY 342 SEITLEALALAGLUASPVALLLEUHIAPMEGLYVALIIEALAMECTHRSERASP 361  
 DB 1027 ACCATTGCGGCTGAAGACACTTTCATGACATGAGGAGTTTCTCATATCACAGTTCTGAC 1086  
 QY 362 SERGLINAMETGLYARGALAGLYGLUVALIIEPROARGTHTTTPGINTHRALAAPLYS 381  
 DB 1087 TCTCAAGCGATGGCGCGTGGTGGAGTTATCTAGAACTTGGCAAAACGCTGACAAA 1146  
 QY 382 ASNLVLYSELUPHEGLYLYSLEUPROGLYUASPVALLYLSAEPANASAPENPHEARGILE 401  
 DB 1147 AACAGAAAGATTTGGCCCGCTTGAAGAAAGAAAGCCATTAACGACACTTACGATC 1206  
 QY 402 LYEARGTYRILESERLYSTYRTHRIEAPNPROALAEUTHRHISGLYVALSERGLUTYR 421  
 DB 1207 AAACGCTACTGTCTAATATCACCATTAACCCAGCATTCGCTCATGGGATTAAGCGAGTAT 1266  
 QY 422 ILGLISERVVALGLUGLUGLYLYSILEALASPLEUVALIETPANPROALAPHEPHE 441  
 DB 1267 GTAGGTTCAGTAGAGGCGAAGAGTGGCTGACTTGATGTGAGAGCCAGCAATTCCTT 1326  
 QY 442 GYVALYSEPROLYSILEVALIIELYSELGLYMETVALIIPHESEGLUMERGLYASP 461  
 DB 1327 GGGGTGAACCAACATGATCATCAAGGCGGATTCATTGGCTTAAGCAAAATGGCGCAT 1386  
 QY 462 SERAENALASERVALPROTHRPROGLINPROVALIYTYRARGGLUMETPHEGLYHIS 481  
 DB 1387 GCGAAGCCCTTCTATCCCTACCCCAACCGGTTATTAACAGAGAAAGTTGTGCTCATCAT 1446  
 QY 482 GLYLYSVALIYSPHEAPTHRSETIETHRPHENVALIATYRGLUASNGLY 501

DB 1447 GGTAAAGCTTAATACGATTCGAAACATCACTTTGTGTCTCAAGCGCTTATGACAAAGCC 1506  
 QY 502 VALYSGLUYSEUGLYLEUGLUARGGLINVALLEUPROVALYLSAENCYARGASNIIE 521  
 DB 1507 ATTAAGAGAAATTAAGACTTGAAGAGCAAGTGTGCGGTAATAATAATTCAGAAATATC 1566  
 QY 522 THRILYLSAPPHELYSPHEAENAPLYSTHRLALYSLIETHRVALAEPPROLYSTHR 541  
 DB 1567 ACTAAAAAGCATGACATTCACACGACACTACTGCTCATGTAAGGTAATCTGAAACT 1626  
 QY 542 PHEGLUVALPHEVALAEPGLYLYSLEUCYETHSERLYSPROTHRSERGINVALPROLEU 561  
 DB 1627 TACCATGTGTGCGAAGGCGAAGAAATTAATCTTTAAACAGCCAAATAAGTAACTTG 1686  
 QY 562 ALAGINARGTYRTHRPHEPHE 568  
 DB 1687 GCGCACTCTTTAGCATTTTC 1707  
 RESULT 9  
 AC34557  
 ID AC34557 standard; DNA, 1710 BP.  
 AC34557;  
 AC 19-JUN-2003 (first entry)  
 XX  
 DT 19-JUN-2003 (first entry)  
 XX  
 DE Prokaryotic essential gene #16214.  
 DE  
 KW Antisense; de; prokaryotic essential gene; cell proliferation;  
 KW drug design; gene.  
 KW  
 OS Helicobacter pylori.  
 OS  
 PN W0200277183-A2.  
 XX  
 PD 03-OCT-2002.  
 XX  
 PF 21-MAR-2002; 2002W0-US009107.  
 XX  
 PR 21-MAR-2001; 2001US-00815242.  
 PR 06-SEP-2001; 2001US-00948993.  
 PR 25-OCT-2001; 2001US-0342923P.  
 PR 08-FEB-2002; 2002US-00072851.  
 PR 06-MAR-2002; 2002US-0362659P.  
 XX  
 PA (ELIT-) ELITRA PHARM INC.  
 XX  
 PI Wang L, Zamudio C, Malone C, Haeselbeck R, Ohlsen KL, Zyskind JW;  
 FI Wall D, Trawick JD, Carr GJ, Yamamoto R, Foreyth RA, Xu HH;  
 XX WPI; 2003-029926/02.  
 DR P-PSDB; ABU30687.  
 XX  
 PT New antisense nucleic acids, useful for identifying proteins or screening  
 PT for homologous nucleic acids required for cellular proliferation to  
 PT isolate candidate molecules for rational drug discovery programs.  
 PS  
 PS Claim 14; SEQ ID NO 22427; 1766pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid comprising any one of  
 CC the 613 antisense sequences given in the specification where expression  
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:  
 CC (1) a vector comprising a promoter operably linked to the nucleic acid  
 CC encoding a polypeptide whose expression is inhibited by the antisense  
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated  
 CC polypeptide or its fragment whose expression is inhibited by the  
 CC antisense nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)

identifying a gene required for cellular proliferation or the biological pathway in which a proliferation-required gene or its gene product lies or a gene on which the test compound that inhibits proliferation of an organism acts; (9) manufacturing an antibiotic; (10) profiling a compound's activity; (11) a culture comprising strains in which the gene product is overexpressed or underexpressed; (12) determining the extent to which each of the strains is present in a culture or collection of strains; or (13) identifying the target of a compound that inhibits the proliferation of an organism. The antisense nucleic acids are useful for identifying proteins or screening for homologous nucleic acids required for cellular proliferation to isolate candidate molecules for rational drug discovery programs, or for screening homologous nucleic acids required for proliferation in cells other than *S. aureus*, *S. typhimurium*, *K. pneumoniae* or *P. aeruginosa*. The present sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 1710 BP; 537 A; 372 C; 384 G; 417 T; 0 U; 0 Other:

Alignment Scores:  
Pred. No.: 1,36e-199 Length: 1710  
Score: 2297.00 Matches: 416  
Percent Similarity: 86.77% Conservative: 76  
Best Local Similarity: 73.37% Mismatches: 75  
Query Match: 76.59% Indels: 0  
DB: 8 Gaps: 0

US-09-904-994b-3 (1-568) x ACA34557 (1-1710)

```
QY 2 LysMetLysLysGlnGlnTyrValaenThrGlyProThrLysGlyAspLysVala 21
DB 7 AAGATTAGCAGAAAGAAATATGTTTCTATGATGCGCTTACAGCCATTAAGTGA 66
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValaGluHisAspLysThrThrGlyGlu 41
DB 67 TTGGGGGATACAGACTGATCGCTGAAGTAAAGATGATGATACCATTTATGGCGA 126
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro 61
DB 127 CTTAAATTCGGTGGCGGTAAACCCCTGAGAGAGGATGAGCCCAATCCCAACCTTAC 186
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspLysThrGlyIle 81
DB 187 AAGAGAAATTTGGATTTAATCATCACTTAACGCTTTAATCGGATTAACCCGATTTAT 246
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGly 101
DB 247 AAGCGGATATTGGTATTAAGATGCAAAATCGCTGCATTGGTAAAGCGGTAA 306
QY 102 AspMetGlnAspGlyAlaSerProHisMetValaGlyValaGlyThrGluAla 121
DB 307 GACATGCAAGATGGCTTAAACCAATCTTAGCGTGGCTCTGACTGAAAGCCTTAGCC 366
QY 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141
DB 367 GGTGAAGGTTTGAATGTAATCTGCTGGTGGTATTGACACACATCCATTTTACCC 426
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValaThrThrMetPheGlyGlyThrGly 161
DB 427 CAACAAATCCCTTACAGCTTTTGCAGACGGGTGAACAACATGATGGTGGCGAATG 486
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHis 181
DB 487 CTGTGATGAGCACTAATCGACTACTATCACTCCAGCAGAAATTTAAATGAAG 546
QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValaGlyPheLeuGlyLysGlyAsnSer 201
DB 547 CTCAGAGCGGCTGAAGATTTCTATGAACCTTAGGTTTCTTGCTTAAAGGTAA 606
QY 202 SerLysLysGlnLeuValaGluGlnValaGluAlaGlyIleGlyPheLysLeuHis 221
DB 548 GCGCAACTCTTATGATTTTC 1707
```

```
DB 607 AACGACGGAGCTTAGCCGATCAATTAAGCTGGTGCATTTGCTTAAATCCACGA 666
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr 241
DB 667 GACTGGGGGCACACCTCTTGGCAATCAATCATGTTGATGTTGACAGCAAAATACGAT 726
QY 242 ValGlnValaCysIleHisThrAspThrValaAsnGluAlaGlyTyrValaAspAsp 261
DB 727 GTGCAAGTCGCTATCCACACAGACACTTTGAATGAAGCCGTTGGGTGAACACATG 786
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyLysHis 281
DB 787 GCAGCTATTGGCCGAGCAGCACTATGACACTTCCACATGGAAGGTGTCGCGCGACAC 846
QY 282 SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrPro 301
DB 847 GCTCTGATATTATTAAGATAGTGTGTGAACAACATTTCTCCGCTTCCACTAACCC 906
QY 302 ThrIleProTyrThrIleAsnThrValaGluHisLeuAspMetLeuMetThrCysHis 321
DB 907 ACTATCCCTTTCATCTGTGATATCAGACAGCAACACATGACATGCTTATGGTGGCAC 966
QY 332 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 341
DB 967 CACTTGATTAAGATTAAGATTAAGATGTTTCAGTTCCGCTGATTCAGGATCCGCCCTCA 1026
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 361
DB 1027 ACCATTTGGCGGTGAAGACACTTGTGATGACATGAGGAGATTTTCTCAATACCCAGCTTGAC 1086
QY 362 SerGlnAlaMetGlyArgAlaGlyLysValaIleProArgThrTrpGlnThrAlaAsp 381
DB 1087 TCTCAAGCTATGAGTGTGTGGGTGAAGTTTACATGAACCTTGGCAACACTGACAA 1146
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 401
DB 1147 AACCAAAAAAATTTGGCGCTTGAAGAGAAAAAGCGATTAAGACAACTTCAGATC 1206
QY 402 LysArgTyrIleSerLysThrIleAsnProAlaLeuThrHisGlyValaSerGluTyr 421
DB 1207 AACCTTCTGTCTTAATATACCATTAACCCAGGATCGCTCAATGGATGACAGATAT 1266
QY 422 IleGlySerValaGluGluGlyLysIleAlaAspLeuValaIleTrpAsnProAlaPhe 441
DB 1267 GTAGTTCGTGAAGATGGGCAAGTGGCTGATTTGATTTGAGATCCGCAATTTCTT 1326
QY 442 GlyValaLysProLysIleValaIleLysGlyGlyMetValaPheSerGluMetGlyAsp 461
DB 1327 GCGGTAAACCCCAACATGATCATCAAGCGGCTTCAATGGCTTGAATGGGTGAC 1386
QY 462 SerAsnAlaSerValProThrProGlnProValaTyrTyrArgGluMetPheGlyHis 481
DB 1387 GCGAAGCTTCTATCCCTTACCCACACACAGTTTATACAGAAATGTTCCCTCATCT 1446
QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValaSerLysValaAlaTyrGluAsn 501
DB 1447 GGTAAAGCAAAATAGAGTCAAAATCACTTTGTGTCAAGCGGCTTATGACAAAGCC 1506
QY 502 ValLysGluLysLeuGlyLysGluArgGlnValaLeuProValaLysAsnCysArgAsn 521
DB 1507 ATTAAGAAAGATTTAGGCTTGAAGACAGATGTTGCCGTAAATAATGCGAAACATC 1566
QY 522 ThrLysLysAspPheLysPheAsnAspLysTrpAlaLysIleThrValaAspProLys 541
DB 1567 ACTAATAAGAAATGATCAATTCACACATACCGCTTCACTTGAAGTCACTTGAAC 1626
QY 542 PheGluAlaPheValaAspGlyLysLeuCysTrpSerLysProThrSerGlnValaPro 561
DB 1627 TACCAATGTTTGTGTGATGCGCAAGATTAATTTTAAACCGCAATTAAGTGAAGCT 1686
QY 562 AlaGlnArgTyrThrPhePhe 568
DB 1687 GCGCAACTCTTATGATTTTC 1707
```





```

Db 1128 GACATGCAAGATGCGCTTAAACAACTTACCGTAGCGTCTCGCTACCTGAAAGCCCTACCC 1187
Qy 122 Gylgluylwettllethralaglylileapserhilethriapheluseerpro 141
Db 1188 GGTAAAGGTTGATGCTAAACGGCTGGTGTATGACACACATCCATTCATTCATCC 1247
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThrGly 161
Db 1248 CAACAAATCCCTACAGCTTTTGCAAGCGGTGAACACCATGATGTTGGTGGAACCGCT 1307
Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHleArgMet 181
Db 1308 CTTCTGATGCGCACTAAAGCACAATCACTACCTCCAGCAGAAAGAAATTTAAATGATG 1367
Qy 182 LeuAlaAlaAlaGluGlyLysSerMetAsnValGlyPheLysGlyLysSerSer 201
Db 1368 CTCAGACCGCGCTGAAGAATATTCTATGAATTTAGGTTCTTGCTAAAGTAAAGCTTCT 1427
Qy 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHleGly 221
Db 1428 AACGATGCGAGCTTAGCCGATCAATGAAACCGGTGGATGGCTTTAAATTCACGAA 1487
Qy 222 AspTrpGlyThrThrProSerAlaIleAspHleCysLeuSerValAlaAspGlyLysTrp 241
Db 1488 GACTGGGGGACACACTCTCTTCGCAATCAATCACTGCTTGAATGTTGGGCAAAATACGAT 1547
Qy 242 ValGlnValCysIleHleThrAspThrValAsnGluAlaGlyLysValAspAspThrLeu 261
Db 1548 GTGCAGATCGCTATCCACACAGACACTTGAATGAAGCGGTGGTGTGAAGACACTATG 1607
Qy 262 AsnAlaMetAsnGlyValAlaIleHleAlaLysThrIleGluGlyAlaGlyGlyHle 281
Db 1608 GCTGCTATGCTGAGCAGCATATGACACACTTCCACACTGAAGCGGTGGGCGGAGAAC 1667
Qy 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301
Db 1668 GCTCTGATATTTAATTAAGTAGCGGTAACACAAATCTTCCGCTTCCACTAACCC 1727
Qy 302 ThrIleProGlyThrIleAsnThrValAlaGluHleAspMetLeuMetThrCysHle 321
Db 1728 ACCATCCCTTACCGTGAATATACAGAAAGACAGACATGACATGCTTATGTTGGTGCAC 1787
Qy 322 HleLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 1788 CACTTGATTAAGACATTAAAGAGATGTTCACTGCTGATTCAGAGATCCGCCCTCAA 1847
Qy 342 SerIleAlaIaGluAspValLeuHleAspMetGlyValIleAlaMetThrSerSerAsp 361
Db 1848 ACCATTCCGCTGAAGACACTTTCATGACATGCGGATTTTCTCAATCACCATGTTCTGAC 1907
Qy 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
Db 1908 TCTCAACCGATGCGCGCTGTGGTGAAGTATCACTGAATCTTGCAAAACGCTGACAA 1967
Qy 382 AsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
Db 1968 AACCAAGAAAGATTTGCGCTTAAAGAAAGAAAGCGGCTAAACGCACTTACAGATC 2027
Qy 402 LysArgGlyIleSerLysTrpThrIleAsnProAlaLeuThrIleGlyValSerGlyLys 421
Db 2028 AAACGCTACTTGTAAATACACATTAACCCAGCATGCTCTATGGATTAAGCAGATAT 2087
Qy 422 IleGlySerValGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
Db 2088 GTAGGTTCAAGTAGAAGTGGGCAAGTGAATGCTGATGTGAGACGACATTCCTT 2147
Qy 442 GlValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461
Db 2148 GGCGTGAACCCCAACATGATATCAAAAGCGGATTCATGCGTTAAAGCAAAATGCGCAT 2207
Qy 462 SerAsnAlaSerValProThrProGlnProValTrpTrpArgGluMetPheGlyHleHle 481

```

```

Db 2208 GCGAAGCGTTCTACTCCATACCCCAACACCGGTTTATTACAGAGAATCTTGCCTCAT 2267
Qy 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTrpGluAsnGly 501
Db 2268 GGTAAAGCTTAATATCGATGACAAACATCATCTTTGTCTCAAGCGGCTTATGACAAAGCC 2327
Qy 502 ValLysGluLysLeuGlyLysGluValGlnValLeuProValLysAsnCysArgAsnIle 521
Db 2328 ATTAAAGAGATTAAGACTTGAAGAACAAGTGTGCGGTAAATAAATTCAGAAATATTC 2387
Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2388 ACTAAAAAGACATGCAATTCACAGACACTACTCTCATTTGAATCAATCTCGAAACT 2447
Qy 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
Db 2448 TACCATGTGTTGCGATGAGCAAGAAGTAATCTTAACACGCAATTAAGTGAGCTTG 2507
Qy 562 AlaGlnArgTrpThrPhePhe 568
Db 2508 GCGCACTCTTACATTTTC 2528

RESULT 12
AAL60579
ID AAL60579 standard; DNA; 1710 BP.
AC AAL60579;
AC AAL60579;
DT 03-SEP-2003 (first entry)
DX Helicobacter pylori urease B (ureB) DNA.
XX Helicobacter pylori; plant pollen; urease B; ureB; ds.
XX Helicobacter pylori.
OS Helicobacter pylori.
PN WO2003044050-A1.
XX 30-MAY-2003.
XX 19-NOV-2002; 2002WO-KR002154.
XX PF 19-NOV-2001; 2001KR-00071712.
XX PA (PARK/) PARK H.
XX PI Park H;
XX DR WPI; 2003-457590/43.
XX PT Producing a recombinant protein using plant pollen, for industrial uses,
XX PT comprises introducing a target gene into an Agrobacterium vector, and
XX PT infecting the transformed Agrobacterium into a cultured plant pollen.
XX PS Example 1; Page 29-30; 34pp; English.
XX CC The invention relates to a method of producing a recombinant protein
XX CC using plant pollen. The method involves introducing a target gene into an
XX CC Agrobacterium vector and infecting the transformed Agrobacterium into a
XX CC cultured plant pollen. The method is useful for producing recombinant
XX CC proteins for testing, diagnosis and prevention and for industrial use.
XX CC The present sequence is Helicobacter pylori urease B (ureB) DNA. This
XX CC sequence is used in the exemplification of the invention
XX SQ Sequence 1710 BP; 540 A; 371 C; 383 G; 416 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.79e-199 Length: 1710
Score: 2291.00 Matches: 415
Percent Similarity: 86.60% Conservative: 76
Best Local Similarity: 73.19% Mismatches: 76
Query Match: 76.39% Indels: 0
DB: Gaps: 0

```

```

US-09-904-994B-3 (1-568) x AAL60579 (1-1710)
QY 2 LysMetLysLysGlnGlnIuYrValaentThrTYrGlyProThrLysGlyAspLysValaIarg 21
Db 7 AAGATTATGACGAGAAAATAATGTTTCTATGATGCGCCCTACTACAGCGCATTAAGAGGAGA 66
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValaGluHisAspTYrThrThrTYrGlyGluGlu 41
Db 67 TTGGGGGATACACACTTGAATCGGTGAAGTACGAAATGATACCACTTATATGGCGAAGAG 126
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
Db 127 CTTAATAATGCGCGTGGTAAACCCCTAAGAGAAAGCATGAGCAACTTAACAACCTTACG 186
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTYrThrGlyIleTYr 81
Db 187 AAGAGAAATCGATCTTAATCATCACTTAACGCTTAAATCGTGATTAACAACCGTATTTAT 246
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHieGlyIleGlyLysAlaGlyAsnLys 101
Db 247 AAGCGGATATTTGGTATTAAGATGCGAAATGCGTGGCATTTGGTAAAGCGGTAAACAA 306
QY 102 AspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeuAla 121
Db 307 GACATGCAAGATGCGCTTAAACAAATCTTAGCGGTGGTCTGCTACTGAGAGCCTTAGCC 366
QY 122 GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
Db 367 GGTGAAGGTTTGAATCTGCTGGTGGTATTGACACACATCTCACTTCACTCTCCCC 426
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValaIleThrMetPheGlyGlyGlyThrGly 161
Db 427 CAACAATAATCCCTACACGCTTTTGCAGCGGTACACACATGATGATGTGGCGGAATCGGC 486
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
Db 487 CCGTGGATGGGCACTTAACCAACCACTATCACTCCAGGTAGAAGAAATTTAAATGAGAG 546
QY 182 LeuArgAlaAlaGlnGluIuYrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 201
Db 547 CTGAGGCGGCAAGAAATAATTTCTATGAATTAAGTTTCTTAGCTTAAGGTAAACGGTTCT 606
QY 202 SerLysLysGlnLeuValGlnGluValaGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221
Db 607 AAGCATGCAAGCTTAGCGCATTAATGAAGCGGTGGCATTTGGCTTAAATCCACGAA 666
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTYrAsp 241
Db 667 GACTGGGGCAGCAGCTCTCTGCAATCAATCAATGAGTGTGCGGACAAATACGAT 726
QY 242 ValGlnValCysIleHisThrAspThrValaAsnGluAlaGlyTYrValaAspAspThrLeu 261
Db 727 GTGCAAGTGGCTATCCACAGACACTTGAATGAAGCGGTGTGTAAAGACACTTAG 786
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTYrHisIleGluGlyValaGlyGlyHis 281
Db 787 GGAGCGCATTTGGCGGAGCAGCATATGACACTTCCACAGAGAGCGCTGTGGCGGACAC 846
QY 282 SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThrPro 301
Db 847 GCTCCCTGATATTAATTAAGTAGCTGTGAACACAAATCTGCGCCCTTCCACTAAGCC 906
QY 302 ThrIleProTYrThrIleAsnThrValaAlaGluHisLeuAspMetLeuMetThrCysHis 321
Db 907 ACTATCTCTTTCAGCTGTGAATACAGAGAGCAAGAACATGATGATGATGTTAGTGTCAC 966
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 967 CACTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1026
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361

```

```

Db 1027 ACTATTGGCGGTGAAGACACTTTGCATGACATGGGGAATTTCTCATGACCAAGTTCTGAC 1086
QY 362 SerGlnAlaMetGlyArgAlaGlyValaIleProArgThrTrpGlnThrAlaAspLys 381
Db 1087 TCTCAAGCTATGGGTGCTGGGTGAAGTATACACAGAACTTGGCAACAGCTGACAA 1146
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
Db 1147 AACAAAAAAGAAATTTGGCGCGCTTGAAGAGAAAGAAAGAGATACAGCAACTTCAGGATC 1206
QY 402 LysArgTYrIleSerLysTYrThrIleAsnProAlaLeuThrHisGlyValSerGluTYr 421
Db 1207 AAACCTCTCTGCTTAATAATACCACTTAACCAAGCATGCGCTCATGAGGATTAAGCAGAT 1266
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValaIleTrpAsnProAlaPhePhe 441
Db 1267 GTAGTTCGTGAAGAGTGGCAAGTGGCTGACTTGTTGTGTGAAGTCCCGCATTTCTTT 1326
QY 442 GlyValLysProLysIleValIleLysGlyLysMetValaIlePheSerGluMetGlyAsp 461
Db 1327 GCGGTGAAGCCCAACATGATCATCAAGCGGATTCATTGCATTGATCAATGGGTAT 1386
QY 462 SerAsnAlaSerValProThrProGlnProValTYrTYrArgGluMetPheGlyHis 481
Db 1387 GCGAAGCTTCTATCTTATCCCAACCAACCGGTTTATTATGAAGAAATGTTCCGTCATCAT 1446
QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValaIleTYrGluAsnGly 501
Db 1447 GGTAAAGCTTAATATAGATGCAACATCACTTTGTGTCTCAAGCGCTTATACAAAGGC 1506
QY 502 ValLysGluLysLeuGlyLysGluArgGluValaLeuProValLysAsnCysArgAsnIle 521
Db 1507 ATTAAGAAGAAATTAAGGCTTGAAGCAAGTGGCCGGTAAATAATGACGAACATC 1566
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValaAspProLysThr 541
Db 1567 ACTAAAAAAGACATGCAATTCACACACACTACCGCTCACTTGAAGTCAATCTGAAACT 1626
QY 542 PheGluValaPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValaProLeu 561
Db 1627 TACCATGTGTTGTGATGCGCAAGAGAAAGTAACTTTAAACCGCAATTAAGTGAAGCTTG 1686
QY 562 AlaGlnArgTYrThrPhePhe 568
Db 1687 GCACAACCTTTAGCATTTTC 1707
RESULT 13
ABA00817
ID ABA00817 standard; cDNA, 1717 BP.
XX
AC ABA00817;
XX
DT 01-APR-2003 (first entry)
XX
DE H. pylori urease B coding sequence.
XX
KW Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;
KW immune response; circulatory vessel; Gut Associated Lymphoid Tissue;
KW GALT; immune response; antigen; Salmonella; Cholera; immunity;
KW Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;
KW hormone; hormone receptor; cancer; ss.
XX
OS Helicobacter pylori.
XX
PN WO200236949-A1.
XX
PD 05-DEC-2002.
XX
PF 24-MAY-2002; 2002MO-AU000661.
XX
PR 25-MAY-2001; 2001AU-00005241.
XX
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

```



XX Mckenzie B, Boyle J, Lew A;  
 XX WPI; 2003-156789/15.  
 XX Raising an immune response in an animal by administering composition  
 PT comprising carrier and antigen bound to targeting moiety which binds to  
 PT receptor present in circulatory vessels in Gut Associated Lymphoid  
 PT tissue.  
 XX  
 PS Disclosure; Page 28-29; 45pp; English.  
 XX  
 CC This sequence encodes H. pylori urease B. This sequence was used in the  
 CC method of the invention for raising an immune response in an animal. The  
 CC method comprises administering to the animal a composition comprising a  
 CC carrier and an antigen bound to a targeting moiety which binds to at  
 CC least one receptor present in circulatory vessels in Gut Associated  
 CC Lymphoid Tissue (GALT). The method is useful for raising an immune  
 CC response in an animal against antigens from Salmonella, Cholera,  
 CC Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut  
 CC associated toxins, gut hormones, gut hormone receptors or gut associated  
 CC cancers. The method is useful for raising both mucosal and systemic  
 CC immunity against any antigen used in the composition  
 XX  
 SQ Sequence 1717 BP; 536 A; 372 C; 388 G; 421 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No: 1,38e-198 Length: 1717  
 Score: 2286.00 Matches: 415  
 Percent Similarity: 86.42% Conservative: 75  
 Best Local Similarity: 73.19% Mismatches: 77  
 Query Match: 76.23% Indels: 0  
 DB: Gaps: 0  
 US-09-904-994B-3 (1-568) x ABA00817 (1-1717)

DB 553 CTGAGGGGCTGAAGATATTCTATGATTTAGGTTCTTGCGCTAAAGGTAAGCTTCT 612  
 QY 202 SerLySvSGuLeuValGluGlnValAlaGlyValAlaIleGlyPheLyLeuHisGlu 221  
 DB 613 AACGATGCGAGCTTAGCCGATCAATTAAGAACCCGGTGGATGGCTTTAAATTCACGAA 672  
 QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTrpAsp 241  
 DB 673 GACTGGGGCACACCTCTTCTGCAATCATCATCGTTAGATGTGGCGGACAAATACGAT 732  
 QY 242 ValGlnValCysIleHisThrAspThrValAlaAsnGluAlaGlyTrpValAspAspThrLeu 261  
 DB 733 GTGCAAGCGCTATCCACAGACACATTTGATAGCCGGTGTGTGAGAACACATATG 792  
 QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTrpHisIleGluGlyAlaGlyGlyHis 281  
 DB 793 GCTGCTATTGGTGGACCATGACATGACATTTCCACATGAGCGGTGGCGGACAC 852  
 QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301  
 DB 853 GCTCCTGATATTATTAAAGTAGCCGGTGAACACAACTTCTCCCGCTTCACATACCCC 912  
 QY 302 ThrIleProTrpThrIleAsnThrValAlaGluHisIleuAspMetLeuMetThrCysHis 321  
 DB 913 ACCATCCCTTTCACCGAATACAGAACAGACATGACATGCTTATGTGTGTCAC 972  
 QY 322 HisLeuAspLyAsArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
 DB 973 CACTTGATTAAGACATTTAAAGAAAGATTGTCAGTTGCTGATTCAGAGATCCGCCCTCAA 1032  
 QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361  
 DB 1033 ACCATTGCGGTGAAGACACTTTCATGATGACATGGGATTTTCCAAACACCACTTCGAC 1092  
 QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLyS 381  
 DB 1093 TCTCAAGCGATGGGCGGTGGGAGTTATCTGAATCTTGAGCAACAGCTGACAA 1152  
 QY 382 AsnLySvSGuLeuPheGlyLyLeuProGluAspGlyLySvAspAsnAspAsnPheArgIle 401  
 DB 1153 AACCAAGAAAGATTTGGCCGTTGAAGAAAGAAAGCGGATTAACGAACTTCAGATC 1212  
 QY 402 LysArgTrpIleSerLySvTrpThrIleAsnProAlaLeuThrHisGlyValSerGluTrp 421  
 DB 1213 AAACGCTACTGTCTTAATATCACCATTAACCCAGCGATCGCTCATGAGCAAGATAT 1272  
 QY 422 IleGlySerValGluGluGlyLySvIleAlaAspLeuValIleTrpAsnProAlaPhePhe 441  
 DB 1273 GTCGTTCTGTAGAAAGTGGCGAAAGTAGCTGACTGGTATTTGTGGAGTCCAGATTCCTT 1332  
 QY 442 GlyValLySvProLySvIleValIleLeuGlyGlyMetValAlaPheSerGluMetGlyAsp 461  
 DB 1333 GCGGTGAACCTTAACATGATCATCAAGAGTGGGTTCATTGCAATTAACCAATGGCGAT 1392  
 QY 462 SerAsnAlaSerValProThrProGlnProValIleTrpArgIleMetPheGlyHisHis 481  
 DB 1393 GCGAAACCTTCTAACCTTACCCCTCAACCGGTTTATTAACAGAAAGTTGCTCATCGT 1452  
 QY 482 GlyValAlaLySvPheAspThrSerIleThrPheValSerIleValAlaTrpGluAsnGly 501  
 DB 1453 GGTAAAGCTTAATACATGCAACATCATCTTTGTGTCTCAAGCGGCTATGACAAAGGC 1512  
 QY 502 ValLySvGluLySvLeuGluArgGlnValLeuProValLySvAsnCysArgAsnIle 521  
 DB 1513 ATTAAAGAAAGATTAAGACTTGAAGACAAAGTGTTCGGGTAAATAATTCAGAAATC 1572  
 QY 522 ThrIleLySvAspPheLySvPheAsnAspLySvThrAlaLySvIleThrValAspProLySvThr 541  
 DB 1573 ACCAAAGAAAGATGCAATTCACAGACATACCGCTCACAATTGAAGTCAATCCGAAACT 1632  
 QY 542 PheGluValPheValAspGlyLySvLeuCysThrSerLySvProThrSerGlnValProLeu 561

DB 1633 TACCATGTGTTGCGATGAGCAAGAAAGTAACCTTTAAACCACTATAAAGTAGCTTG 1692  
QY 562 AAGAGTATGTTTThrPhephe 568  
DB 1693 GCGCAACTCTTAGCATTTTTC 1713  
RESULT 14  
ABA00816  
ID ABA00816 standard; cDNA; 1719 BP.  
AC ABA00816;  
DT 01-APR-2003 (first entry)  
DE H. felis urease B coding sequence.  
XX  
XX Gene; rat; IgG2a; light chain; E. coli; galactosidase; immunogen;  
KW Immune response; circulatory vessel; Gut Associated lymphoid tissue;  
KW GALT; immune response; antigen; Salmonella; Cholera; immunity;  
KW Helicobacter pylori; HIV; Candida; P. gingivalis; gut; parasite; toxin;  
KW hormone; hormone receptor; cancer; ss.  
XX Helicobacter felis.  
OS  
PN W0200296949-A1.  
XX  
XX 05-DEC-2002.  
XX  
XX 24-MAY-2002; 2002MO-AU000661.  
XX  
XX 25-MAY-2001; 2001AU-00005241.  
PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.  
PI McKenzie B, Boyle J, Lew A;  
XX  
XX WPI; 2003-156789/15.  
XX  
XX Raising an immune response in an animal by administering composition  
PT comprising carrier and antigen bound to targeting moiety which binds to  
PT receptor present in circulatory vessels in Gut Associated lymphoid  
PT tissue.  
XX  
XX Disclosure; Page 27-28; 45pp; English.  
XX  
XX This sequence encodes H. felis urease B. This sequence was used in the  
CC method of the invention for raising an immune response in an animal. The  
CC method comprises administering to the animal a composition comprising a  
CC carrier and an antigen bound to a targeting moiety which binds to at  
CC least one receptor present in circulatory vessels in Gut Associated  
CC lymphoid tissue (GALT). The method is useful for raising an immune  
CC response in an animal against antigens from Salmonella, Cholera,  
CC Helicobacter pylori, HIV, Candida, P. gingivalis, gut parasites, gut  
CC associated toxins, gut hormones, gut hormone receptors or gut associated  
CC cancers. The method is useful for raising both mucosal and systemic  
CC immunity against any antigen used in the composition  
XX  
SO Sequence 1719 BP; 490 A; 411 C; 421 G; 397 T; 0 U; 0 Other;  
Alignment Scores:  
Pred. No.: 4.94e-196 Length: 1719  
Score: 2258.00 Matches: 414  
Percent Similarity: 85.54% Conservative: 71  
Best Local Similarity: 73.02% Mismatches: 82  
Query Match: 75.29% Indels: 0  
DB: 8 Gaps: 0  
US-09-904-994B-3 (1-568) x ABA00816 (1-1719)  
QY 2 LysMetLysLysGlnGluTyrValAsnThrTyrGlyProThrIlySerGlyAspLysValArg 21  
DB 15 AAGATTTCAGAAAGAAATATGTTTCTATGTATGTGTCCTACACTACCGGGAGTCGTGTTAGA 74

QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrThrGlyGluGlu 41  
DB 75 CTCGGCGACACTGATTTGATCTTAGAAGTGAGCATGATGTGACACACTTAGTGTAAGAG 134  
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61  
DB 135 ATCAAAATTTGGGGCGGTAACCTATCCGTGATGGATGATCAACCAATAGCCCTTAC 194  
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeuAspTyrThrGlyIleTyr 81  
DB 195 TCTTATGAATTAATGATTTGGTGTCTCACTAACGCCCTCATGTGATGACTATAGGCAATTTAC 254  
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101  
DB 255 AAAGCCGACATTGGGATTTAAAGAGCAAGATTGACAGCATTTGGCAAGCCAGCAATAAG 314  
QY 102 AspMetGlnAspGlyValSerProHisMetValGlyValGlyThrGluAlaLeuAla 121  
DB 315 GACATGCAAGATGGCGCTAATATATCTTTCGTAGGTCCTGTCTACAGAGGCTTTGGCA 374  
QY 122 GlyGluGlyMetIleLeuThrAlaGlyIleLeuAspSerHisThrHisPheLeuSerPro 141  
DB 375 GCTGAGGCGCTGATTTGTAACCGCTGTGACATGATGACGATATTCATTTCTCTCC 434  
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThrGly 161  
DB 435 CAACAATATCCCTACTGCTTTTCCAGCGGGTTTACACCATGATTTGAGAGGACAGAGA 494  
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181  
DB 495 CCGCGGATGGGACGAAATCGAACACCATCATCTCCGCGGCTAATATTAAGAACTAAG 554  
QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201  
DB 555 TTGCGTGACGCCGAGAAATACCGCATTAATCTTAGGCTTTTGCTTAGGGAGATGTCT 614  
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221  
DB 615 TACGAACCTCTTTACCGCATGATGAAAGAGGGGCGATTGGTTTAAATCCACGAA 674  
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241  
DB 675 GACTGGGAGAACACACTGCTATTCACACTGCTCAATGTCGCCCATGATTCGAT 734  
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261  
DB 735 GTGCAAGTGCTATCCACACCGATACCTTAACAGGGCGGCTGTATAGAAACACCTTA 794  
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValaGlyGlyHis 281  
DB 795 GAGGCGATTGCCGGGGGACCATCCATCTTCACACTGAAGGGGCTGGGGGTGGACAC 854  
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301  
DB 855 GCTCCAGATGTTATCAAAATGCGAGGGAAATTAACATTTCAACCCGCTCTCAACCCG 914  
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321  
DB 915 ACCATTCCTTTTCAACCAAAACCTGAAGCCGAGACATGACATGTATATGTGTGCCAC 974  
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
DB 975 CACTTGATTAAGATTAAGAAAGATGCGAGTTGTCGATTCGAGATTCGCCCCCAA 1034  
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361  
DB 1035 ACTATGCGGCTGAAGACCACTCATGACATGGGATCTTTTCTATCCACGCTCCGAC 1094  
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381  
DB 1095 TCTCAGGCTTATGGAGCGCTTAGCGGTGATACACGCACTTTGGCAGACAGACGAA 1154  
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401

```

Db 1155 AACAAAAAGGTTGGCGCTTGAAGAGAAAAAGCGCATACGACCACTCCGCATC 1214
Qy 402 LysArgTyrIleSerIleThrIleAsnProAlaLeuThrIleGlyValSerGluTyr 421
Db 1215 AAAGCGTACATCTCTAAATACACCATCAACCCCGCATCGGCGATGGATTTCTGACAT 1274
Qy 422 IleGlySerValIleGluGlyIleValIleAspLeuValIleTPAsnProAlaPhephe 441
Db 1275 GTGGGCTCTGTGGAGAGGGGCAATACGCCGACCTCGTCTTGGAGTCCGGCTTTCTT 1334
Qy 442 GlyValIleProIleValIleValIleGlyValIleGlyValIlePheSerGluMetGlyAsp 461
Db 1335 GGCACTTAAGCCCAATATGATATTATTAAGGCGGATTTATGCGCTCTCTCAATGGCGAT 1394
Qy 462 SerAsnIleSerValIleProIleProGlnProValIleTyrTyrGluMetPheGlyHisHis 481
Db 1395 GCCAATCGCTATTCCACCCCTCAGCCCTCTATTACCTGAAATGTTTGGACACCAT 1454
Qy 482 GlyValIleAspPheAspThrSerIleThrPheValSerIleValIleTyrGluAsnGly 501
Db 1455 GGGAAAAACAAATTCGACACCAATATCATCTTCTGTTCCAAAGCGGCTTACAGCGAGG 1514
Qy 502 ValIleGluIleValIleGluIleValIleValIleProValIleAsnGlyArgAsnIle 521
Db 1515 ATCAAAAGAGAACTAGGCGTATGATCGCGTGTATTTGCCAGTGAAGAACTGTCCGCAATTC 1574
Qy 522 ThrIleValIleAspPheIlePheAsnAspIleThrIleValIleThrValIleProIleThr 541
Db 1575 ACTAAAAAGGACCTCAAAATTCACGATGTGACCCACATATTGATGTCACCCCTGAAACC 1634
Qy 542 PheGluValIlePheValIleAspGlyIleValIleValIleSerIleValIleProIleValIle 561
Db 1635 TATAAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1694
Qy 562 AlaGlnArgTyrThrPhePhe 568
Db 1695 GCGCACTTATATATTGTTTC 1715

RESULT 15
ID AAQ12485 standard; DNA, 2767 BP.
AC AAQ12485;
XX
DT 27-AUG-2003 (revised)
DT 23-SEP-1991 (first entry)
DE DNA encoding A and B subunits of H. pylori urease.
XX
KM Gastritis; peptic ulceration; duodenitis; helicobacter; campylobacter; ss.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT CDS 64..717
FT FT /*tag= a
FT FT /label= subunit A
FT CDS 721..2400
FT FT /*tag= b
FT FT /label= subunit B
FT FT /note= "31 kD"
XX
PN WO9109049-A.
PD 27-JUN-1991.
XX
PF 19-DEC-1989; 89GB-00028625.
PR 19-DEC-1989; 89GB-00028625.
XX
PA (THRE-) 31 RES EXPL LTD.

```

```

XX
PI Tabagchal IS, Claydon CL, Wren BW, Kleanthous H;
XX
DR WPI; 1991-208084/28.
XX
DR P-PSDB; AAR12515, AAR13550.
XX
PT Oligo:nucleotide(s) specific for Helicobacter pylori - used as probes and
PT primers to detect H pylori infection, in diagnosis of gastritis, and
PT duodenal and peptic ulceration.
XX
PS Disclosure; Page 16; 28pp; English.
XX
CC The DNA is a 2.7 kb TagI fragment encoding the A and B subunits of H.
CC pylori (previously C. pylori) urease, i.e. the 66 and 31 kD antigens.
CC From the sequence, probes and primers can be designed for the
CC amplification (by PCR) of the gene, to produce a probe. Common to all H.
CC pylori strains so far tested and which do not occur in other ureases and
CC can therefore be used as a specific indication of the presence of H.
CC pylori. H. pylori (previously C. pylori) is strongly implicated in the
CC pathogenesis of gastritis and duodenal and peptic ulceration in man. The
CC primers/probes can be used for the detection of H. pylori DNA in gastric
CC mucosa, saliva or faecal samples to provide an early diagnosis of
CC infection. See also AAQ12434-Q12486. (Updated on 27-AUG-2003 to correct
CC OS field.)
XX
SQ Sequence 2767 BP; 867 A; 550 C; 635 G; 715 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5,01e-195 Length: 2767
Score: 2250.00 Matches: 412
Percent Similarity: 85.71% Conservative: 74
Best Local Similarity: 72.66% Mismatches: 81
Query Match: 75.03% Indels: 1
DB: 2 Gaps: 0

US-09-904-994B-3 (1-568) x AAQ12485 (1-2767)
Qy 2 LysMetIleValIleGluIleValIleAsnThrIleThrIleProIleValIleValIleValIle 21
Db 790 AAGATTAGCAAGAAAGATAGCTTATGATAGGCGCTTACAGCGCATTAAGTGAAGA 849
Qy 22 LeuGlyAspThrAspLeuThrAlaGluValIleHisAspTyrThrThrTyrGlyGluGlu 41
Db 850 TTGGCGCATACAGCTTATGATCGCTGAGTAGAATACATACATCACTTATGCGAAGAG 909
Qy 42 LeuIlePheGlyValIleGlyIleThrIleArgGluGlyIleMetGlyIleSerAsnSerProAsp 61
Db 910 CTTAAATTCCGCGCGCGGTAAACCTTAAGAGAGGATGAGCCATCTAACACCCCTAGC 969
Qy 62 GluAsnThrIleAsnLeuValIleThrAsnAlaMetIleIleAspTyrThrThrIleTyr 81
Db 970 AAAGAAAGACTGATCTTAATCATCATCACTTAACGCTTAACTGATTAACCGGATTTAT 1029
Qy 82 LysAlaAspIleGlyIleValIleValIleValIleHisGlyIleGlyValIleValIleValIle 101
Db 1030 AAAGCGGATATTGATATTAAAGATGCAAAATCGCTGCGATTTGTAAGCGGTAACAA 1089
Qy 102 AspMetGlnAspGlyValIleSerProIleMetValIleValIleValIleThrGluAlaLeuAla 121
Db 1090 GACACGCAAGATGGCGTTAAAAACAACTTAAGCGTGGCTCTGCTACTGAAGCCTTAGCC 1149
Qy 122 GlyGluIleMetIleIleThrAlaGlyIleValIleAspSerIleThrIleAspLeuSerPro 141
Db 1150 GGTAAAGTTGATGTAACTGCTGCTGATTAAGACACACATCCACTCATCTCCCC 1209
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrIleThrIlePheGlyIleGlyThrGly 161
Db 1210 CAACAAATCCCTACAGCTTTTGCAGCGGTGAACAAACATGATGTGCGGAACTGGC 1269
Qy 162 ProValIleGlyIleThrAsnAlaThrThrIleThrIleProGlyIleTyrPheAsnLeuHisArgMet 181
Db 1270 CTGCTGATGCACTAACGACCAACCACTATCATCTCAGGTAGAGAAATTTAAATTCATG 1329

```

QY 162 LeuArgAlaIaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201  
|||||  
Db 1330 CTCAGAGCGCGTGAAGAATATCTATGAACCTTGCTTCTGCTTAAGTAAACGCTTCT 1389  
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaIleGlyPheLysLeuHISGlu 221  
:::  
Db 1390 AACGATGCAGAGCTTAGCCGATCAAAATGAAGCTGGCGCTTAGCTTAAATCCACGAA 1449  
QY 222 AAPTTPGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241  
|||||  
Db 1450 GACTGGGGCACCACTCTTCTGCAATCAATCAATGCTTGATGCTTGCGGACAAATACGAT 1509  
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261  
|||||  
Db 1510 GTGCAAGTGGCTATCCACACAGACACTTGAATGAAGCCGGTGGCTGAAGACACTATG 1569  
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaIleHisIleGluGlyAlaGlyLysHis 281  
:::  
Db 1570 GCAGCTATTGGCCGGACGCACTATGCACACTTACACACAGGCGCTGGCGGACAC 1629  
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301  
:::  
Db 1630 GCTCCTGATATATTAATTAAGTGGCGGTGAACACACATCTCTCCGCTTCACTAACCC 1689  
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHISLeuAspMetLeuMetThrCysHis 321  
|||||  
Db 1690 ACTATCCCTTTCACCGCTGAATACAGAAAGCCGAACACATGCATGCTTATGGTGTCCAC 1749  
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
|||||  
Db 1750 CACTTGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTA 1809  
QY 342 SerIleAlaIaGluAspValLeuHISAspMetGlyValIleAlaMetThrSerSerAsp 361  
:::  
Db 1810 ACCATTTGGCGGTGAAGACACTTGCATGCATGCGGATTTTCTCAATCACTAGTTCTGAC 1869  
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381  
|||||  
Db 1870 TCTCAAGCGATGGCGGTGTGGTGAAGTTATCATAGAACTTGGCAACAGCTGACAAA 1929  
QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401  
|||||  
Db 1930 AATTAATAAAGAAATTTGGCGCTTGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1989  
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGluTyr 421  
|||||  
Db 1990 AAACGCTACTTGTCTTAATTAACCATTAACCCAGCGATCGCTCATGGGATTAAGCGAGTAT 2049  
QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhePhe 441  
:::  
Db 2050 GTCGGTTCTGTGAAGTGGCAAGTGGCTGACTTGTGTATGTGGAGTCCCGCATTTCTTT 2109  
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461  
|||||  
Db 2110 GGTGTGAAGCCCAACATGATCAAAAGGCGGTTTCATCATGATGCAAAATGGGTGAT 2169  
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHisHis 481  
:::  
Db 2170 GCGAAGCTTCTATCTCTTACCCCAACACAGCTTATTAACAGAGAAATGTTCCGCTCATCAT 2229  
QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501  
|||||  
Db 2230 GGTAAAGCTTAATACATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGGC 2289  
QY 502 ValLysGluLysLeuGlyLeuGluLysGlnValLeuProValLysAsnCysArgAsnIle 521  
:::  
Db 2290 AATTAAGAAAGAAATTTGGGCTTAAAGACAAAGTGTGCGGTAAATAATTGCAAGAAATATC 2349  
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541  
|||||  
Db 2350 ACTAATAAAGACATGCAATTCACGACACTACCGCTCACATGAACTCAATTCGAAACT 2409  
QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561

Db 2410 TACCATGTGTTCTGTGATGGCAAGA-AGTAACTTAAACGCAATTAAGTGAAGCTTG 2468  
QY 562 AlaGlnArgTyrThrPhePhe 568  
Db 2469 GCGCAACTCTTATGACATTTTC 2489

Search completed: February 15, 2005, 21:02:10  
Job time : 703.99 secs

RESULT 1					
BD185302					
LOCUS	BD185302	2883 bp	DNA	linear	PAT. 17-JUN-2003
DEFINITION	Helicobacter felis vaccine.				
ACCESSION	BD185302				
VERSION	BD185302.1	GI:1877502			
KEYWORDS	JP 2002355054-A/1.				
SOURCE	Helicobacter felis				
ORGANISM	Helicobacter felis				
	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;				
	Helicobacteraceae; Helicobacter.				
REFERENCE	1 (bases 1 to 2883)				
AUTHORS	Kusters,J.G. and Cattoli,G.				
TITLE	Helicobacter felis vaccine				
JOURNAL	Patent: JP 2002355054-A 1 10-DEC-2002;				
	AKZO NOBEL NV				
COMMENT	OS Helicobacter felis				



DEFINITION Sequence 1 from Patent EP1176192.  
ACCESSION AX356683  
VERSION AX356683.1 GI:18674020  
KEYWORDS  
SOURCE Helicobacter felis  
ORGANISM Helicobacter felis  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Helicobacteraceae; Helicobacter.  
REFERENCE  
1  
Kusters, J.G. and Cattoil, G.  
Helicobacter felis vaccine  
Patent: EP 1176192-A 1 30-JAN-2002;  
Akzo Nobel N.V. (NL)  
FEATURES  
Source Location/Qualifiers  
1..2883  
/organism="Helicobacter felis"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:214"  
206..886  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAD3255.1"  
/db\_xref="GI:18674021"  
/translation="MKLTPKOEKELLYAGEVARKRKAREGLKINPEALAYSAHIM  
DEARRGKTYAQLMEBCMHFLKDEWNPFGNNVPDQVEATPPDGRKLTVMNPPIE  
DEHFKAGEVFGCDKDELNAKEVTELEVTGSPKSLHVGSHFHEANAKLFDEE  
KAYGKRLDIPSGNLTLRIGAQTRAVOLIPLGSKKVIQMGVLVNNIADERHKHALDK  
AKSHGFIR"  
897..2603  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=1  
/protein\_id="CAD3256.1"  
/db\_xref="GI:18674022"  
/translation="MKMKQBYVNTYPTGPKDKVRLGDTDLMAVEHDYTYGBELKE  
GAKTIREMGOSNPEDNTLDVITAMIIDYGIKADIGIKNGKIHGKXGND  
MONGVSPHMYGCTELALAGEKMITTNGDISHTFISPOOFALANGVTMMGGST  
GPDVGNATTTIDFKMNLHRMLRAEYSIMNVGFLGKNSSKQLVEQVAGAIIGK  
LHEDWGTTPSAIDHCLSVADYDVQCIHNDVNEADYVDDTLNANMGRALHAHIGS  
AGGSHSPDIVTMAGELNLPSSPTPIPTINTVAEHLDMLTCHMDKRIREDLQPS  
OSRIRPESIAEVDLHDMGVYAMTSSDSQMGRAGEVYPTWQADNKKKEFLPDK  
GKNDNFRIRKRYISKYTINPALTHGVEYGVSEKGLADLVVANNPAPFGVKEPLVLC  
GGMVPSREMGDSNASVPTPPVYTRFMFGHGRKAKPTSTITPYSKVAIENGVAKGL  
ERQVLPEVNCRNITKDFKFNDAKITVDPKTEFVVDGKLCTSKPTSOVPLAQRYT  
F"

ORIGIN  
Alignment Scores:  
Pred. No.: 3.59e-190 Length: 2883  
Score: 2999.00 Matches: 568  
Percent Similarity: 100.00% Conserved: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0  
US-09-904-994B-3 (1-568) x AX356683 (1-2883)

QY 1 MetLysMetLysLysGInGluTYrValAsnThrTYrGlyProThrLysGlyAspLysVal 20  
Db ATGAAATGAAAAACAAGAATATGTAATATCCACGACCCACCAAGGCGATTAAGTG 956  
QY 21 ArgLysGlyAspThrAspLeuTTPAlaGluValGluIuhIAspTYrThrTYrGlyGlu 40  
Db CGCTTAGAGATACCGATCTTTGGCGAGAAAGTAGAAGATGACTATACCACTATGGCGAA 1016  
QY 41 GluLeuLysPheGlyAlaGlyLysThrLLeaGlyGlyMetGlyGlnSerAsnSerPro 60  
Db GAACTTAAATTTGGCGGGGAAAAACTATCCGTAGAGGTATGGGTACAGAGAAATAGCCTT 1076  
QY 61 AspGluAsnThrLeuAspLeuValLLeaThrAsnAlaMetLLeIleAspTYrThrGlyLe 80  
Db GATGAAACACCCCTTAGATTAAGTATCATACCGCATGATTAATGACTACACCGGATTT 1136

QY 81 TyrLysAlaAspLLeGlyYlLeLysAsnGlyLysLLeIAspGlyYlLeGlyYlLeGlyYlLe 100  
Db TACAAAGCCGACATTTGGATTAATAAAACGCAAAATTCATGGCATTTGCCAAGCGGAAC 1196  
QY 101 LysAspMetGlnAspGlyValSerProHLeaMetValaGlyValaGlyThrGluAlaLeu 120  
Db AAGGACATGCAAGATGGCGTGAAGCTTCATATGGTCTGTTGGGTGGGACAGAAACCTA 1256  
QY 121 AlaGlyGlyMetLLeIethrAlaGlyGlyLLeaSerHLeaThrHLeaPheLeuSer 140  
Db GAGGGAGAGATATGATTAATACCGTGGGGGATTCATTCACACCACTTCTTCTTCT 1316  
QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValaThrMetPheGlyGlyYlThr 160  
Db CCACAAATTCCTTACCGCTTACGCAAAAGGGGTTAACCAATGTTTGAAGGGGACA 1376  
QY 161 GlyProValAspGlyThrAsnAlaThrThrLLeaProGlyLysTPAsnLeuHLeaArg 180  
Db GGTCCTTAGATGGACGAATGCGACTATCATCTCCGGGCAATGAACTTGACACCGC 1436  
QY 181 MetLeuArgAlaAlaGlyGlyTYrSerMetAsnValaGlyPheLeuGlyLysLysSer 200  
Db ATGTGGCGGCGAGAGAAAGTATTTATGAATGGGCTTTTGGCCAAAGGCAATAGC 1496  
QY 201 SerSerLysLysGlnLeuValaGlnValaGlnAlaGlyAlaLLeGlyPheLysLeuHLe 220  
Db TCTAGCAAAACAACTTGTAAGAACAGTGAAGCGGCGCGATGTTTAAATTCAT 1556  
QY 221 GluAspTPGlyThrThrProSerAlaLLeaPheLysSerValaLLeaAspGlyTyr 240  
Db GAACTTGGGGCACACACCAAGTGCATGCATCATCTTGAACCTTGGCGAGATGATAC 1616  
QY 241 AspValaGlnValaCysLLeaIethrAspThrValaAsnGluAlaGlyTYrValaAspAspThr 260  
Db GATGTCAAGTTGTATCCACACCGATACAGTCAATGAGGAGGATTAATGATGACACC 1676  
QY 261 LeuAsnAlaMetLysGlyAlaGlyAlaLLeaAlaLLeaIethrAlaGlyYlLeaGlyYl 280  
Db CTAAATGCAATGAACGGGCGCGCTTCACCTCAACATTTGAGGAGCGGGGAGGA 1736  
QY 281 HisSerProAspValaLLeaThrMetAlaGlyLysLeuAsnLLeaPProSerSerThr 300  
Db CACTCACCTATGTTATCACCATGCGAGCGAGCTCAATTTCTACCTCTCTCCACACC 1796  
QY 301 ProThrLLeaProTYrThrLLeaAsnThrValaLLeaIuhIAspMetLysMetThrCys 320  
Db CCCACTATTCCTTATACATTAATATAGGTTGAGAAACATTAAGCATGATGACATGC 1856  
QY 321 HisHisLeuAspLysArgLLeaArgLysAspLeuGlnPheSerGlnSerArgLLeaArgPro 340  
Db CACCACCTTACAAACCAATCCCGGAGGATTTCAATTTCTCAAGCCCTATCCGCC 1916  
QY 341 GlySerLLeaLLeaGlyAspValaLLeuHLeaAspMetGlyValaLLeaLLeaMetThrSerSer 360  
Db GGCTTATCGCGGCTGAAGATGCTCCATGATGAGGTGATGCGGATGACAGCTCG 1976  
QY 361 AspSerGlnAlaMetGlyArgAlaGlyGluValaLLeaProArgThrTPGlnThrAlaAsp 380  
Db GATTCGCAACCAATGGGCGGTGAGCGGAGAGTATCTCGAATCTTGACGAGATCGGAT 2036  
QY 381 LysAsnLysLysGlyPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400  
Db AAGGATTAATAAAAGATTTGGTAAGCTTCCGAAGATGCGAAAGATTAACGTAATTTCCG 2096  
QY 401 LLeaLysArgTYrLLeaSerLysTYrThrLLeaAspProAlaLeuThrHLeaGlyValaSerGlu 420  
Db ATTBAAGGCTTACATCTCAATATCATCTATCAACCCCGCTTGAACCGCGTGAAGGAG 2156  
QY 421 TyrLLeaGlySerValaGlyGlyLysLLeaLLeaAspLeuValaLLeaAsnProAlaPhe 440  
Db TATATCGGCTCTGTGGAAGAGGCGCAAGATGCGCACTTGTGTGTGATTCCTGCTTT 2216

Qy	441	PheGlyValIyGProIySileValIleIyGlyGlyMetValIaPheSerGlyMetGly	460
Db	2217	TTTGGCGTAAACCCAAATCGATCAAAAGGGGATGTGTCTTCTCTGAAATGGGC	2276
Qy	461	AAPSerAsnAlaSerValProThrProGlnProValTyrTyrArgIyMetPheGlyHis	480
Db	2277	GATTCTAAACGCGCTGTGCCACTCCCAACCGGTTATTATACCGCAAAATGTTGGGCAAT	2336
Qy	481	HisGlyIyValAlaIySHePheSerThrSerIleThrPheValSerIySValAlaTyrGlyuan	500
Db	2337	CACGGCAAGGCGAAATTTGACACACATCACTTTGTTCCAAAGTCGCTATGAAAAAT	2396
Qy	501	GlyValIyGlyIySLeuGlyLeuGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIy	520
Db	2397	GGCGTAAAGAAAGAGCTGGGCTTGAAGCCGCAAGTTCTACCGGTCAAAAACCTCGTAAC	2456
Qy	521	IleThrIyAlaIySAspPheIySPhaeAsnAspIySThrAlaIySIIeThrValAspProIyS	540
Db	2457	ATCACCAAGAAAGACTTCAAGTTCAACGACAAACGGCAAAATCACCCGTCATCCGAAA	2516
Qy	541	ThrPheGlyIyValPheValAspGlyIySLeuGlySThrSerIySProThrSerGlnValPro	560
Db	2517	ACCTTGAGAGTCTTTGTAGATGGCAAACTCTGCACCTTAACCACTCGCAAGTGCGT	2576
Qy	561	LeuAlaGlnArgTyrThrPhePhe	568
Db	2577	CTAGCCCAAGCGCTACACTTTCTTC	2600
RESULT 3			
BD185306		2452 bp	DNA
LOCUS	BD185306		Helicobacter felis vaccine.
DEFINITION	BD185306		
ACCESSION	BD185306.1		GI:31877506
VERSION	JP 2002355054-A/5.		
KEYWORDS	Helicobacter felis		
SOURCE	Helicobacter felis		
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.		
REFERENCE	1	(baaes 1 to 2452)	
AUTHORS	Kusters,J.G. and Cattooli,G.		
TITLE	Helicobacter felis vaccine		
JOURNAL	Patent: JP 2002355054-A 5 10-DEC-2002;		
COMMENT	AKZO NOBEL, NV		
OS	Helicobacter felis		
PN	JP 2002355054-A/5		
PD	10-DEC-2002		
PF	16-JUL-2001 JP 2001214711		
PR	17-JUL-2000 EP 00202565, 8		
PI	JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOOLI		
PC			
C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC			
A61K39/23,			
PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,			
PC C12N1/19,			
PC			
C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/			
53,			
PC			
G01N33/53, G01N33/566, G01N33/569//C12N9/80, C12R1:01, C12Q1/68, PC			
C12R1:01,			
PC C12N15/00, C12N5/00, A61K37/02			
CC Helicobacter felis vaccine			
PH Key	Location/Qualifiers		
FT CDS	(48) . . (728)		
FT CDS	(739) . . (2445) .		
FEATURES			
source	1. .2452		
	/organism="Helicobacter felis"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:214"		
ORIGIN			
+			

Alignment Scores:			
Pred. No.:	1,586-188	Length:	2452
Score:	2973.00	Matches:	562
Percent Similarity:	99.65%	Conservative:	4
Best Local Similarity:	98.94%	Mismatches:	2
Query Match:	99.13%	Indels:	0
DB:	6	Gaps:	0
US-09-904-994B-3 (1-568) x BD185306 (1-2452)			
Qy	1	MetIySmetIySLeuGlyGlnIyValIleThrIyGlyProThrIyGlyAspIySVal	20
Db	739	ATGAAAAATGAAAAAACAAGATATGTAATATACCTACCGGCCACACACAGCGCATTAAGTG	798
Qy	21	ArgLeuGlyAspThrAspLeuThrAlaGlyIyGlyIySAspTyrThrThrIyGlyGly	40
Db	799	CGCTTAGAGATACCGATCTTTGGGACAGAGTGAACATGATTAATACCACTTAATGCGAA	858
Qy	41	GlyLeuIySPhaeGlyAlaGlyIySThrIleArgIyGlyMetGlyGlnSerAsnSerPro	60
Db	859	GAACTCAAATTCGGTGCAGGTAAACTATCCGTAGGGTATGGTCCAGAGCAATAGCCA	918
Qy	61	AspGlyAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle	80
Db	919	GATGAAAAACCTTAGATTAGTATGATACCAACGCGAGATATTATGATACACCGGAT	978
Qy	81	TyrIySAlaAspIleGlyIleIySAsnGlyIySIIeIyGlyIleGlyIySAlaGlyAsn	100
Db	979	TACAAAGCCGACATTTGGCTTAAATGCAAAATTCATGSGCATTTGGCAAGCGAGAAC	1038
Qy	101	IySAspMetGlnAspGlyValSerProHisMetValIleGlyValIleGlyThrGlyIyAlaIy	120
Db	1039	MAGCAATGCAAGATGGCGTAAAGCCCTCATATGTCGTGGTGGGCAACAAGACATA	1098
Qy	121	AlaGlyIyGlyMetIleIleThrAlaGlyIyIleAspSerHisThrHisPheLeuSer	140
Db	1099	GCAGGGGAAGGTATATTATTCCTGGGGGATGATTCACACACCACTTCTCTCT	1158
Qy	141	ProGlnIySProThrAlaIySAsnGlyValIleThrMetPheGlyIyGlyIyThr	160
Db	1159	CCACAACATTCCTTACCGCTTACGCAATGGCGTTACAACATGTTGGCGGTGACATA	1218
Qy	161	GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyIySLeuIySAsnIySArg	180
Db	1219	GGCCCGTAGAGGACAGATCGACTACATCTCCGGGCAAAATGAACTTGACCGGC	1278
Qy	181	MetLeuArgAlaIleGlyIySLeuIySAsnValIySLeuGlyIySAsnSer	200
Db	1279	ATGTTGCGGCGACAGAAAGATATTCATGAATGTGGGCTTTTGGCAAAAGCAATAGC	1338
Qy	201	SerSerIySLeuGlyIySLeuValIleGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIy	220
Db	1339	TCTAGTAAAAACAATTGTACACAGTGAAGAGGGGGCATTTGTTAAATTTGCAT	1398
Qy	221	GlyAspTyrGlyIyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyIy	240
Db	1399	GAAAGCTGGGGCAACAATCCAGTCGATCATCTGTTGAGGTGAGATGAAATAC	1458
Qy	241	AspValIyGlnValCysIleHisThrAspThrValAsnGlyIyGlyIyValAspAspThr	260
Db	1459	GATGTGCAAGTTGTATACACCCGATACGGGTACATGAGGCGGTATATGATATACACC	1518
Qy	261	LeuAsnAlaMetAsnGlyIyArgAlaIleHisAlaTyrHisIleGlyIyAlaIyGlyIyGly	280
Db	1519	CTAAATGCAATGAAGGGGGCGCATCCATGCTACACATTTGAAGGAGCGGATGAGGA	1578
Qy	281	HisSerProAspValIleThrMetAlaGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIyGlyIy	300
Db	1579	CACCTACCGATGATATACCACTGACAGGCGAAGATATTCTACCTCCCTCCCAACAC	1638
Qy	301	ProThrIleProTyrThrIleAsnThrValAlaGlyIySLeuAspMetLeuMetThrCys	320
Db	1639	CTTACTATCCCTATACATTAATACGTTGAGAACTTAGACATGCTTATGACCTGC	1698



QY 321 H i h i s t e u a p l y s a r g i l l e a r g l u a s p l e u g l i n p h e s e r g l i n s e r a r g i l l e a r g p r o 340  
 D b 1699 C A C C A C C T A G A T T A A A C G A T C C G A G A G A T T C C A A T T T T T C C A A A G C G T A T C C G C C C 1758  
 QY 341 G l y S e r i l l e a l a g l u a p v a l l e u h i s a p m e t g l y a l l e a l a m e t t h r s e r 360  
 D b 1759 G G C T C T A T C C G C G T G A A G A T G T C T C C A T A T A T C G T G A T C G C A T A C A A G T T C C 1818  
 QY 361 A s p S e r g l a m a m e t g l y a r g l a g l y l u a l l e p r o a r g t h r t r p g l i n t h r a l a a s p 380  
 D b 1819 G A T T C G A C A G C A A T G G G G C G C T G G G A A G T G A T T C C T A A C T T G C C A A C T G C A G A C 1878  
 QY 381 L y s e n t l y s l y s g l u p h e g l y l y l e u p r o g l u a s p g l y l y s a p a s n a s p a s n p h e a r g 400  
 D b 1879 A A G A A T A A A A A A A A T T G T G A C C T T C C G A A A G T G A G A T A A G A C A C A C T T C C G C 1938  
 QY 401 I l e l y s a r g l y r i l e s e r l y s t y r t h r l e a n p r o a l e u t h r i l e g l y a l s e r g l u 420  
 D b 1939 A T C A A A C G T A T A T C T C C A A A T A C A C C A T T A T C C C G T T G A C C C A T G C G T A G C A G 1998  
 QY 421 T y r i l l e g l y s e r v a l g l u g l u g l y l y s i l e a l a s p l e u v a l l a t r p a s n p r o a l a p h e 440  
 D b 1999 T A T A T C G C T C T G T G A A G A G G C A A G A T C G C G A C T T G G T G T G G A A T C T G C C T T T 2058  
 QY 441 P h e g l y v a l l y s p r o l y s i l e a l l e l y g l y g l y m e t v a l l a p h e s e r g l u m e r g l y 460  
 D b 2059 T T T G C C T A A A C C C A A A A T G T A T C A A A G C G T A T G T G T G T C T C T G A A A T G G C C 2118  
 QY 461 A s p S e r a n a l a s e r v a l p r o t h r p r o g l i n p r o v a l t y r t y r a r g l u m e t p h e g l y h i s 480  
 D b 2119 G A T T C T A T A T G C T G T G C C A C T C T C A G C C G T T A T T A C C C G A A A T T T T G G C A T 2178  
 QY 481 H i e g l y l y s a l a l y s p h e a p t h r s e r i l e t h r p h e v a l s e r l y s a l a l a t y r g l u a s n 500  
 D b 2179 C A C G C A A G C G A A A T T T G A C A C C A G A T C A C T T T G T T C C A A A G C G C T A T G A A A A T 2238  
 QY 501 G l y a l l y s g l u y l y s l e u g l y l e u g l u a r g l u a l l e u p r o v a l l y s a n c y a r g a s n 520  
 D b 2239 G G T G T A A A A A A A A C T A G G T T T A G A C G C A A G G T G C C C C G G A A A A C T G C C G T A A C 2298  
 QY 521 I l e t h r l y s a p p h e l y s p h e a s n a p l y s t h r a l a l y s i l e t h r v a l a s p p r o l y s 540  
 D b 2299 A T C A C C A A G A A G A C T T C A A G T T C A A G C A A A A C T G C A A A A T C A C C G T G A T C C G A A 2358  
 QY 541 T h r p h e g l u a l p h e v a l a s p g l y l y l e u c y a t h r s e r l y s p r o t h r s e r g l u a l p r o 560  
 D b 2359 A C C T T C A G A G C T T T G T A G A T G C A A A C T G C A C C T T A A C C A C C T G A A G T G C C T 2418  
 QY 561 L e u a l a g l i n a r g t y r t h r p h e p h e 568  
 D b 2419 C T A G C C A A C G T A C A C T T T C T T C 2442

RESULT 4  
 AX356695  
 LOCUS AX356695 2452 bp DNA linear PAT 13-FEB-2002  
 DEFINITION Sequence 13 from Patent EP1176192.  
 ACCESSION AX356695  
 VERSION AX356695.1 GI:18674032  
 KEYWORDS  
 SOURCE Helicobacter felis  
 ORGANISM Helicobacter felis  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
 Helicobacteraceae; Helicobacter.

REFERENCE  
 1 Kusters, J.G. and Cattoi, G.  
 Helicobacter felis vaccine  
 Patent: EP 1176192-A 13 30-JAN-2002;  
 Akzo Nobel N.V. (NL)  
 Location/Qualifiers  
 1..2452  
 /organism="Helicobacter felis"  
 /mol\_type="unassigned DNA"

CDS  
 /db\_xref="taxon:214"  
 48..728  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAD3263.1"  
 /db\_xref="GI:18674033"  
 /translation="MKLTPEKQEKFLTYAGEVARRRKAGSLKLNQEPALATYSNIM  
 DEARRGKTVAELMBECMFLKQDEVMPGVGNVPDGVATPEPDGTLVTVMPIR  
 DEHFKAGEVFGCDKDIELNAGEVELEVTNGNPKSLAHGSHFPAKALKFEDE  
 KAYGKRDIPISGNTLRIGAGQTRKVQLIPLGSSKKVIGMGLVNNIADERSHKALDK  
 AKSHGFIK"  
 739..2445  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAD3264.1"  
 /db\_xref="GI:18674034"  
 /translation="MKMKQEVNTYGPPTGDKVRLGDDPLMAEVEHDYTYGRELK  
 GAKTIREGSGNSPDENTLDVITTAAMIDVTGYIKADIGIKNGTIRGKRGND  
 MQGVSPHNVGVGTEALAGEGMITAGGIDSHFLSPQFPALANGVTFMFGGT  
 GPVDGTAATTTPGKMLHRLMAEYSNMVQFLGKGNSSKKQLEQVBAHIGK  
 LHEDWGTTPSALIDCLSVADYDVQCIHTDYNAGYVDDTANMGAHIAHYIG  
 AGGHSFDVITMAGEVNIILPSSTPTPIPTINVAEHLDMTCHLDKRLDLPDS  
 OSRIIRPGSLAEDVLDIGVIAMTSSDSQAMGSAEVIPTWTQADKNKKEPGLP  
 GADNDRIRIRYISKITINPALTHGUSEYISYBEGKLADLVYNNPAFPGVKRIVK  
 GAVVESEMDSNASVPTPOPVYREMFHGRKAPDSTIFVSQVAEYGVKGLG  
 ERKVLVPCNCRNITKDFKFNDRKATITVDPKTFEVRVDDKLTCKTSKPSVPLAQRVT  
 F"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,58e-188 Length: 2452  
 Score: 2973.00 Matches: 562  
 Percent Similarity: 99.65% Conservative: 4  
 Best Local Similarity: 98.94% Mismatches: 2  
 Query Match: 99.13% Indels: 0  
 DB: Gaps: 0

US-09-904-994B-3 (1-568) x AX356695 (1-2452)

QY 1 M e t l y s m e t l y s l y s g l u g l y r v a l a s n t h r t y r g l y p r o t h r l y s g l y a s p l y s v a l 20  
 D b 739 A T G A A A A T G A A A A A A C A A G A G T A T G T A A A T A C T A C G A C C A C C A C A G C G A T A A A G T G 798  
 QY 21 A r g l e u g l y a s p t h r a s p l e u t r p a l a g l u a l g l u h i s a p t y r t h r t y r g l y g l u 40  
 D b 799 C G C T T A G A G A T A C C A T C T T T G G C G A G A G T G A A C A T A C T A T A C A C C T A T A G C G A A 858  
 QY 41 G l u l e u l y s p h e g l y a l a g l y l y t h r l l a r g l u g l y m e t g l y g l n s e r a s n s e r p r o 60  
 D b 859 G A A C T C A A A T T C G T G C A G G T A A A A C T A T C C G A G G G T A T G G T C A G A C A A T A G C C C A 918  
 QY 61 A s p g l u a s n t h r l e u a p l e u v a l l e t h r a s n a l a m e t l l e t h r a p p t y r t h r g l y l e 80  
 D b 919 G A T G A A A A C C T T A G A T T A G A T C A C C A G C G A T G A T T A T G A C T A C A C C G G A A T T 978  
 QY 81 T y r l y s a l a s p i l l e g l y l l e l y s a n g l y l y s i l e i s g l y l l e g l y l y s a l a g l y a s n 100  
 D b 979 T A C A A A C C G A C A C T T G C A T T A A A A T G C A A A T C A T A C A T G C A T T G C A A G C A G A A A C 1038  
 QY 101 L y s a s p m e t g l u a s p g l y l a l s e r p r o h i m e t v a l a l g l y a l g l y t h r g l u a l l e u 120  
 D b 1039 A A G A A C A T G C A A A T G C G T A A C C C C A T A T G T C T G T G T G G G C A C A G A A G A C A C T A 1098  
 QY 121 A l a g l y l u g l y m e t l l e i l e t h r a l a g l y g l y l l e a s p s e r h i s t h r i s p h e l e u s e r 140  
 D b 1099 G C A G G G A A G T A T G A T T A T T A C C G C T G S G G A T G A T T C A C A C A C C A C T T C T C T C T 1158  
 QY 141 P r o g l i n p h e p r o t h r a l a l e u a l a s n g l y v a l t h r t h r m e t p h e g l y g l y t h r 160  
 D b 1159 C C A C A C A A T T C C C T A C C G C T A G C C A A T G C G T T A C A C A A T G T T G C G C G T G C A C A 1218

```

Qy 161 GYProValAspGlyThrAsnAlaThrThrIleThrProGlyLeuThrAsnLeuHisArg 180
Db 1219 GGGCCCGGTGATGAGCGCAAGATCGACTACATCACTCCGGCAAAATGAACTTGACACCC 1278
Qy 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200
Db 1279 ATGTTGCGCGAGCAGAAAGATTTCTATGAATGTGGCTTTTGGGCAAAAGCATATAC 1338
Qy 201 SerSerLeuLeuGluLeuValGluGluValAlaGluAlaIleGlyPheLeuHis 220
Db 1339 TCTAGTAAATAAACAACTTGTAGAACAGTAGAAGCGGGCGCATGTGGTTTAAATTGCAT 1398
Qy 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyr 240
Db 1399 GAAAGACTGGGCGCAACCTCCAACTGCGATCGATCACTGCTTGACCTTACAGATGAATAC 1458
Qy 241 AspValGluValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260
Db 1459 GATGTGCAAGTTTGTATACACACCGATACGGTCAATGAAGGAGTTATGTAGTACACAC 1518
Qy 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGluValAlaGlyGly 280
Db 1519 CTAAATGCAATAGAACCGGCGCATCGCTTACACAAATGAGGAGCGGTTGAGGA 1578
Qy 281 HisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr 300
Db 1579 CACTCACCTGATGTTATATACCATGGAGCGAAAGTAATTTCTACCTCCTCCACAAAC 1638
Qy 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320
Db 1639 CCTACTATCCCTATACCATTAATGAGTTGCGAAGACCTTAGACATGCTTATGACCTGC 1698
Qy 321 HisHisLeuAspLysArgIleArgGluAspLeuGluPheSerGlnSerArgIleArgPro 340
Db 1699 CACCACTAGATTAAGCAATCGCGAGATCTCCAAATTTCTCAAGCGGATCCGCCCC 1758
Qy 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360
Db 1759 GGCTCTATCGCGCGTGAATGTCTCCATGATATCGGTGTATCGCGATGACAAATTTCC 1818
Qy 361 AspSerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAsp 380
Db 1819 GATTCGCAAGCAATGGGCGCGCTGGGGAAGTATCTTAGAACTTGAGCAAACTCGACAGAC 1878
Qy 381 LysAsnLysLeuGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAspAsp 400
Db 1879 AAGAAATTAATAAGAAATTTGTAAGCTTCCGAAAGATGTGCAGATATAGCAAACTTCCGC 1938
Qy 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420
Db 1939 ATCAAAACGCTATATCTCCAAATACCACTTAATCCGCTTTGACCAATGGCGTAGCGAG 1998
Qy 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValAlaTrpAsnProAlaPhe 440
Db 1999 TATATCGCTCTGTGGAAAGGCGCAAGATCGCCGACTGTGTGGAATCCTGCTTT 2058
Qy 441 PheGluValLysProLysIleValIleLysGlyGluMetValValPheSerGluMetGly 460
Db 2059 TTTGGCGTAAACCCAAATTCGATCAAGCGGATGATGATGTGTTCTCGAAATGGGC 2118
Qy 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480
Db 2119 GATTCGTAAGCGCTGTGGCACTCTCAAGCGGTTTATTAACGGCAAAATGTTTGGCAT 2178
Qy 481 HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn 500
Db 2179 CACGGCAAGCGCAAAATTTGACACACACATCACTTTGTTCCAAAGTCCCTATGAAAAAT 2238
Qy 501 GlyValLysGluLysLeuGlyLeuGluValArgGluValIleLeuLysAsnCysAspAsn 520
Db 2239 GGTGTCAAAAGAAAACTTAGGTTTAGAGCGCAAGGTCTCCCGTGAAAAACGCGCTAAC 2298
Qy 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys 540

```

```

Db 2299 ATCACCAAGAGAGACTTCAGATTCAACGCAAAACTGCAAAAATCACCCTGCATCGAAA 2358
Qy 541 ThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnAlaPro 560
Db 2359 ACCTTCGAGGCTTTGTGTAGATGGAACCTGTGACCTTAAACCACTTGAAAGTGCTT 2418
Qy 561 LeuAlaGlnArgTyrThrPhePhe 568
Db 2419 CTAGCCCAACGCTACACTTTCTTC 2442

RESULT 5
BD185303 2405 bp DNA linear PAT 17-JUN-2003
LOCUS Helicobacter felis vaccine.
DEFINITION
ACCESSION BD185303.1 GI:31877503
VERSION BD185303.1 GI:31877503
KEYWORDS JP 2002355054-A/2.
SOURCE Helicobacter felis
ORGANISM Helicobacter felis
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
Helicobacteraceae; Helicobacter.
REFERENCE
1 (bases 1 to 2405)
Kusters,J.G. and Catoili,G.
Helicobacter felis vaccine
Patent: JP 2002355054-A 2 10-DEC-2002;
AKZO NOBEL NV
OS Helicobacter felis
PN JP 2002355054-A/2
PD 10-DEC-2002
PF 16-JUL-2001 JP 2001214711
PR 17-JUL-2000 EP 00202565.8
PI JOHANNES GERARDUS KUSTERS,GIOVANNI CATTOLO
PC
C12N15/09,A61K38/00,A61K39/106,A61K39/118,A61K39/12,A61K39/175, PC
A61K39/23,
PC A61K39/235,A61K39/39,A61K39/395,A61P1/04,A61P31/04,C12N1/15,
PC C12N1/19,
PC C12N1/21,C12N5/10,C12N9/80,C12Q1/68,G01N33/15,G01N33/50,G01N33/
53,
PC
G01N33/53,G01N33/566,G01N33/569//C12N9/80,C12R1/01,C12Q1/68, PC
C12R1/01,
PC C12N15/00,C12N5/00,A61K37/02
CC Helicobacter felis vaccine
FH Key Location/Qualifiers
FT CDS (1)..(681)
FT CDS (692)..(2398)
FEATURES
source location/Qualifiers
1..2405
/organism="Helicobacter felis"
/mol_type="genomic DNA"
/db_xref="taxon:214"
ORIGIN
Alignment Scores:
Pred. No.: 1,13e-187 Length: 2405
Score: 2960.00 Matches: 560
Percent Similarity: 99.30% Conservative: 4
Best Local Similarity: 98.59% Mismatches: 4
Query Match: 98.70% Indels: 0
Gaps: 0
DB: 6
US-09-904-994B-3 (1-568) x BD185303 (1-2405)
Qy 1 MetLysMetLysLeuGluGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20
Db 692 ATGAAATAAGAAAAACAAGATATGTAACACTTACGAGCCACCAAGCGATATAAGTG 751
Qy 21 ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrTyrGlyGlu 40
Db 752 CGCTTAGAGATACCGATCTTTGGGCAAGAAAGTAACAATGACTATACCACTTATGGCGAA 811

```

Qy 41 GluLeuLeuSpheGlyValAglyLeuThrIleArgGluGlyMetGlyGlnSerHisSerPro 60  
 Db 812 GAGCTCAAAATTTGGCGGGTAATAAATCCGTAGGGGTATGGGCAAGCAAGTAAGGCCA 871  
 Qy 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetCileIleAspTyrThrGlyIle 80  
 Db 872 GATGAAAAACCTTATGATTTAGTATGATCCAAACCGGATGATTTGACCTACACCGGGATT 931  
 Qy 81 TyrIleValAspIleGlyIleLeuAsnGlyLeuIleHisGlyIleGlyLeuAlaGlyAsn 100  
 Db 932 TATTAACCCGCACTTGGTATTAATAATGCAAAATCCATGGTATTTGGCAAGCGGGGAAAC 991  
 Qy 101 LeuAspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu 120  
 Db 992 AAAACACATGCAAGATGGCGTAAAGCCCTCATATGTCGGGTGTGGGCAAGAACACTA 1051  
 Qy 121 AlaGlyGluGlyMetCileIleThrAlaGlyIleAspSerHisSerHisPheLeuSer 140  
 Db 1052 GCAGGGAAAGGTATGATTTATTAACCGCTGGGGGATCGATTGGCACACCCACTTCTCT 1111  
 Qy 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr 160  
 Db 1112 CCCCAACATTCCTACCGCTCTAGCCCATGCTTACCAACATGTTTGGAGGTGGACA 1171  
 Qy 161 GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyGlyThrPheAsnLeuHisArg 180  
 Db 1172 GGTCCGGTAGATGGCAGAAATGCAACCACTCATCTCCGGGCAATGCACTTGACCGC 1231  
 Qy 181 MetLeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLeuGlyAsnSer 200  
 Db 1232 ATGTGGCGCGCATGTAAGATATTCTATGATGTTGGGCTTTTGGGCAAGGCAATTGC 1291  
 Qy 201 SerSerLeuGlyGlnLeuValGluGlnValGluAlaGlyAlaGlyPheLeuLeuHis 220  
 Db 1292 TCCAGTAAAAAACAACCTGTAAGAACAGTAGAAGCGGGCGGATGGCTTTAAATTTGCAT 1351  
 Qy 221 GluAspTrpGlyThrThrProSerAlaIleAspHisGlyLeuSerValAlaAspGluTyr 240  
 Db 1352 GAAGACTGGGGCACAACAACCAAGTGCATGATCATGCTTGACCGTGAAGTAATAC 1411  
 Qy 241 AspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr 260  
 Db 1412 GATGTGCAAGTTTGTATCCACACCGATACCGTCAATGAGGAGTTATGTAGACACCC 1471  
 Qy 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyIle 280  
 Db 1472 CTAAATGCGATGAACGGGGCGGCATCATGCTTACCAATGAGGGAGCGGAGAGA 1531  
 Qy 281 HisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr 300  
 Db 1532 CACTGACCTATGTTATCACCATGCGAGCGCATCATATTTTCACTCTCCACACACC 1591  
 Qy 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320  
 Db 1592 CCCACTATTCCTTATACCATTTAATACGTTTCAGAAACCTTAGACATGCTATGACCTGC 1651  
 Qy 321 HisIleLeuAspLeuArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340  
 Db 1652 CACCACTTAGACAAACGATCCGCGAGATCTCCAGTTTCCAAAGCCGATCCGCCCC 1711  
 Qy 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360  
 Db 1712 GGCCTATTTCCCTGTAAGATGTGCTCATATATTTGGCGGATCGCGATACAAAGCTCG 1771  
 Qy 361 AspSerGlnAlaMetGlyArgAlaGlyValIleProArgThrTrpGlnThrAlaAsp 380  
 Db 1772 GATTTCGAAACATGGGGCGGCTGGGAAAGTATTCCTTAAGACTTGGCAAACTGCCAGAC 1831  
 Qy 381 LeuAsnLeuValGlyGluPheGlyLeuLeuProGluAspGlyLeuAspAsnAspAsnPheArg 400  
 Db 1832 AAGAAATAAAAAAGAAATTTGGTAAGCTTCTGTAAGATGTTGACATTAATGACAACTTCGC 1891

Qy 401 IleAspArgTyrIleSerIleTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420  
 Db 1992 ATCAACGCTATATCTCCAAATACCACTTAATCCCGCTTTGACCATGGGTAGCGAG 1951  
 Qy 421 TyrIleGlySerValGluGluGlyLeuIleAlaAspLeuValIleTrpAsnProAlaPhe 440  
 Db 1952 TATATCCGCTCTGTGGAAAGAGGCAAGATCCGCCACTTGGTGTGTGAAATCTCTTTC 2011  
 Qy 441 PheGlyValIleProValIleValIleLeuGlyGlyMetValValPheSerGluMetGly 460  
 Db 2012 TTTGGTGTAAACCAAAATGTGATCAAAAGCGGTATGTTGGTGTCTTGAATAGGCG 2071  
 Qy 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480  
 Db 2072 GATTCTAACCGGTGTGGCCACACTCAGCGGGTTATTAACCGAAATGTTGGGCAT 2131  
 Qy 481 HisGlyLeuAlaValPheAspThrSerIleThrPheValSerLeuValAlaTyrGluAsn 500  
 Db 2132 CACGGCAAGGGCAATTTGACACACGATCACTTTGTTCCAAAGTCGCTTATGAAAT 2191  
 Qy 501 GlyValIleGlyGluLeuGlyLeuGluArgGlnValLeuProValIleAsnCysArgAsn 520  
 Db 2192 GCGGTAAAGAAAACTAGCTTAGCGCAAGGTGCTACCCGTGAAAACTGCCGCAAC 2251  
 Qy 521 IleThrIleValAspPheLeuPheAsnAspLeuThrAlaValIleThrValAspProLeu 540  
 Db 2252 ATCACTAGAAAGCTTCAATTCACAAACAAAGCGCGCATACATGTCGATCTTAAA 2311  
 Qy 541 ThrPheGluValPheValAspGlyLeuGlyCysThrSerLeuProThrSerGluValPro 560  
 Db 2312 ACCTTCAGAGCTTTTGTAGATGCGAAACTCTGCACTTAAACCCGCTCTGAAGTGCT 2371  
 Qy 561 LeuAlaGluArgTyrThrPhePhe 568  
 Db 2372 CTAGCCCAAGCTACACTTCTTC 2395

RESULT 6  
 AX356686  
 LOCUS AX356686 2405 bp DNA linear PAT 13-FEB-2002  
 DEFINITION Sequence 4 from Patent EP1176192.  
 ACCESSION AX356686  
 VERSION AX356686.1 GI:18674023  
 KEYWORDS  
 SOURCE Helicobacter felis  
 ORGANISM Helicobacter felis  
 Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 Helicobacteraceae; Helicobacter.

REFERENCE  
 1 Kusters, J.G. and Catoili, G.  
 Helicobacter felis vaccine  
 Patent: EP 1176192-A 4 30-JAN-2002;  
 Akzo Nobel N.V. (NL)

FEATURES  
 source  
 1..2405  
 Location/Qualifiers  
 /organism="Helicobacter felis"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:214"  
 1..681  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAD3257.1"  
 /db\_xref="GI:18674024"  
 /translation="MKLTPEQKFLLYAGVARKRKGLKLNQPEALAYISAHIM  
 DEARRGKTVASLMEBCMHPLKKQBYVPGVNVPLGVATPPDGTKLVTVMPIEP  
 DEHRKGEVFGDDKDIELNAGKEVLELTNPGPSLHNGSHHPETKALKFDDE  
 KAYGKRLDIPSGNTLRIAGQTRKNQILPLGSKKYI GNMGLVNNIADENHKKALKDK  
 AKSHGRFK"  
 692..2398  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAD3258.1"

CDS

/db xref="GI:18674025"  
 /translation="MKMKQKQEVNVTYPTTGDVKVRLGDTDLMAVEHDVYTYGEEIKF  
 GAKGKTRBEKGSNSPDENTLDLVITNMIIDYTGYYKADIGIKNGKHGIGKAKND  
 MODGVSPHNVGVTEALAGEBNITTAGIDSHTEPLSQQPTALANKVTFPGGGT  
 GVDGDTNATITTPGKNLHRMLRAAEYMNVPFLGKSSSKQLVEVEGALIFK  
 LHEDGTPSAIDHCLSVADYDVOVCIHTDVNEGYVDLTMANGALAHYHIEG  
 AGGSPDIVITAGELNLPSSSTTPTPTVTVAEHLDMTCHLDRILEDLOFS  
 OGRIRPGSLAADVLHDIGVIMTSSDSQAMGRAGEVIRPTQTDKNKKEKGPED  
 GADNDFRKIRYISKYTNIPALTHGVSEKIDLVYVMPAPFGVYKVIK  
 GGVVFPSEKGDNSAVPTPQPTYYRBMFGHGAKPDSITTVSKVAIENGYKEKGL  
 EKKVLVKNCRNITKKDFKFNKTAHIITVDPKFEVVDGKLCITSKPASEVPLAORT  
 FE"

## ORIGIN

## Alignment Scores:

Pred. No.:	1,13e-187	Length:	2405
Score:	2960.00	Matches:	560
Percent Similarity:	99.30%	Conservative:	4
Best Local Similarity:	98.59%	Mismatches:	0
Query Match:	98.70%	Indels:	0
DB:	6	Gaps:	0

US-09-904-994b-3 (1-568) x AX356686 (1-2405)

QY 1 MetLysMetLysLeuGluGluValLeuThrTyrGlyProThrLysGluAspLysVal 20  
 DB 692 ATGAAAATGAAAAACAAGATATGTAAACACTACGACCCACCAAGGCAATAAAG 751  
 QY 21 ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGlu 40  
 DB 752 CCTTATGGAGATACCATCTTTGGGCGAAGTAGAACATGACATATCCACTTAATGCGAA 811  
 QY 41 GluLeuLysPheGluValAlaGlyLysThrIleArgGluGluMetGlyValIleSerLysSerPro 60  
 DB 812 GAGCTCAAAATTTGGCGGTAACCTATCCGAGAGGATAGGGGTAGAGCAATATCCCA 871  
 QY 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80  
 DB 872 GATGAAAACACCTTAAATTTAGATGATCAACAAGCGATGATTAATGACATACCGGAAAT 931  
 QY 81 TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100  
 DB 932 TATAAAGCCGACATTTGGTTTAAATAATGCAAAATCATATGATATGCGAAAGCGGGAAC 991  
 QY 101 LysAsnMetGluAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu 120  
 DB 992 AAGAATCATGACATGAGCGGATGCTCATATGCTGAGGTGGGCGACAGAGCACTA 1051  
 QY 121 AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140  
 DB 1052 GCAGGGGAAAGATATGATTTATCCGCTGGGGGATGATTCGACACCCACTTCCCTCTCT 1111  
 QY 141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyIleThr 160  
 DB 1112 CCCCAACAATTCCTCTCCGCTCAGCAATGCGTATCAACCAATGTTGGAGGTGCACAA 1171  
 QY 161 GlyProValAspGlyLysThrAsnAlaThrTrpIleThrProGlyLysThrAsnLeuHisArg 180  
 DB 1172 GGTCCGGATGATGCGCAATGCGACCATCATCCCGGCAAAAGAACTTGACCCGC 1231  
 QY 181 MetLeuArgAlaAlaGluGluLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200  
 DB 1232 ATGTTCCGGCGACCTGAAGATTTCTATGAATGTTGGGCTTTTGGGCAAAAGCAATAG 1291  
 QY 201 SerSerLysLysGluLeuValGluGluValGluAlaGlyAlaIleGlyPheLysLeuHis 220  
 DB 1292 TCCAGTAAAAAACAACCTCGTAGAACAGTAGAAGCCGGCGCGATTGCTTAAATTCGAT 1351  
 QY 221 GluAspTyrGlyLysThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr 240  
 DB 1352 GAAGACTGGGGCAACAACAGTGGATGATCATCTGCTTGAGCGGTAGAGATGAAATAC 1411  
 QY 241 AspValGluValCysIleHisThrAspThrValAsnGluValAlaGlyLysValAspAspThr 260

DB 1412 GATGTCGAAGTTTGTATCCACACCGATACGGTCATATGAGGCGATTATGATGATGACACC 1471  
 QY 261 LeuAsnAlaMetAsnGlyAlaArgAlaIleHisAlaIleTyrHisIleGluGlyValAlaGlyIleGly 280  
 DB 1472 CTAAATGGATGATGAACGGGGCGCCCTCAATGCTTCCACTTGAAGAGCGGGCGAGGA 1531  
 QY 281 HisSerProAspValIleIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThr 300  
 DB 1532 CACTCACTGATGTTATATCAATGCGACAGGAGCTCAATATTCTACCTCCCTCCACACC 1591  
 QY 301 ProThrIleProTyrThrIleAsnThrValAlaGluHisLysAspMetLeuMetThrCys 320  
 DB 1592 CCCACTATTCCTTATACCATTTATACGTTGTGAGAAACATTATGATGATGCTCATGACCTGC 1651  
 QY 321 HisHisLysLeuAspLysAlaGlyIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340  
 DB 1652 CACCACTAGACAAACGCAATCCGCGAGAGATCTCCAGTTTCCCAAGCCGTATCCGCCCC 1711  
 QY 341 GlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360  
 DB 1712 GGCCTTATTTGCCGCTGAAAGATGCTCTCAATATTTGGCGTATGCGCATGACCAAGCTCG 1771  
 QY 361 AspSerGlnAlaMetGlyAlaArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAsp 380  
 DB 1772 GATTGCAAGCAATGGGGCGCGCTGGGGAAGATTCCTCAGAACTTGGCAACTGCAGAC 1831  
 QY 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400  
 DB 1832 AAGAATTAATAAAGAAATTTGGTAGCTTCTGAAAGATGGGAGAAATGACAACTTCCGC 1891  
 QY 401 IleLysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlu 420  
 DB 1892 ATCAAAACGCTATATCTCCAAATACACATTAATCCCGCTTACCCATGCGCTGAGCGAG 1951  
 QY 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValIleTrpAsnProAlaPhe 440  
 DB 1952 TATATCGGCTGTGTGAAGAGGGCAAGATCCGCCGACTTGTGTGTGGAATCTGCTTTC 2011  
 QY 441 PheGlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGly 460  
 DB 2012 TTTGGTGTAAACCCCAAAATCGTATCAAAAGCGGTATGCTGTCTCTTAAATAGGGC 2071  
 QY 461 AspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHis 480  
 DB 2072 GATTCTAACGCTCTGTGCGCACACTCAGCCGGTTTATACCGGAATGTTGGGCAAT 2131  
 QY 481 HisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsn 500  
 DB 2132 CACGGCAAGCGAAATTTGACACACGACATCACTTTGTTCCAAAGTCGCCCTATGAAAT 2191  
 QY 501 GlyValLysGluLysLeuGlyLysLeuArgGluValIleLeuProValLysAsnGlyAsnArg 520  
 DB 2192 GCGGTGAAGAAACCTAGGCTTGAAGCGCAAGGTGTACCCGTAAATAACGCGCAAC 2251  
 QY 521 IleThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLys 540  
 DB 2252 ATCATTAAGAAAGACTTCAAAATTCACAAACAGAGCGGCATATCACTGTGATCTTAA 2311  
 QY 541 ThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValPro 560  
 DB 2312 ACCTTCGAGGCTTTGTGATGAGCAAACTGTGCACCTCAAAACCGCGCTGAAAGTGCT 2371  
 QY 561 LeuAlaGlnArgTyrThrPhePhe 568  
 DB 2372 CTAGCCCAACGCTACACTTCTTC 2395

RESULT 7  
 LOCUS BD185305 2407 bp DNA linear PART 17-JUN-2003  
 DEFINITION Helicobacter felis vaccine.  
 ACCESSION BD185305  
 VERSION BD185305.1 GI:31877505

KEYWORDS JP 2002355054-A/4.  
SOURCE Helicobacter felis  
ORGANISM Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Bacteri; Proteobacteria; Helicobacter.  
REFERENCE 1 (bases 1 to 2407)  
AUTHORS Kusters, J.G. and Cattoil, G.  
TITLE Helicobacter felis vaccine  
JOURNAL Patent: JP 2002355054-A 4 10-DEC-2002;  
AKZO NOBEL NV  
COMMENT OS Helicobacter felis  
PN JP 2002355054-A/4  
PD 10-DEC-2002  
PF 16-JUL-2001 JP 2001214711  
PR 17-JUL-2000 EP 00202565.8  
PI JOHANNES GERARDUS KUSTERS, GIOVANNI CATTOIL  
PC C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC  
A61K39/23,  
PC A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,  
PC C12N1/19,  
PC C12N1/21, C12N5/10, C12N9/80, C12Q1/68, G01N33/15, G01N33/50, G01N33/ PC  
53,  
PC G01N33/53, G01N33/566, G01N33/569// (C12N9/80, C12R1/01), (C12Q1/68, PC  
C12R1/01),  
PC C12N15/00, C12N5/00, A61K37/02  
CC Helicobacter felis vaccine  
FH Key Location/Qualifiers  
FT CDS (2). (682)  
FT CDS (693). (2399).  
FT CDS Location/Qualifiers  
1. .2407  
/organism="Helicobacter felis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:214"

ORIGIN

Alignment Scores:

Pred. No.: 3,31e-187 Length: 2407  
Score: 2953.00 Matches: 558  
Percent Similarity: 99.12% Conservative: 5  
Best Local Similarity: 98.24% Mismatches: 5  
Query Match: 98.47% Indels: 0  
DB: Gaps: 0

US-09-904-994b-3 (1-568) x BD185305 (1-2407)

QY 1 Mettlyemettlysblysgingluytrvaalasnthrrtyrgluyprothrlysglyaspllysva 20  
DB 693 ATGAAATGAAAAACAGAGATGTAATACTACGACCTACCAACGCGACAAAGTG 752

QY 21 Argleuglyasplrrasplleutrpalaglvalglunhsasplrrthrrthrrtyrglvglu 40  
DB 753 CGCTTAGAGATACCGATCTTTGGCGCAGAGTAAACATGACTATACCATTAAGCGCA 812

QY 41 Gluleulysphegiyalaglylvethrrtlearglulgymetgylginserasnsrpro 60  
DB 813 GAGCTCAATTTGGCGCGGTAAACATATCCGTAGGCGATGGGTCAAGACATATGTCCA 872

QY 61 Aspglunserthrleuasplleuvalillethrasmalametlelleasprtyrrthrrtyr 80  
DB 873 GATGAAACACCCCTAGATTAGTATCATCAACACGAGATGATTATGACTACACCGGATTT 932

QY 81 Tyrlysalasplilegilyllelyasnglylysiilehisiagilyllegiyllysalaglyasn 100  
DB 933 TACAAACCCGACATTTGGCATTTAAATAATGCAAAATCCATGSCATTTGGCAAGCAAGAAC 992

QY 101 Lysasplmetglinaaspllyvalserprohismetvalyvaliglyvalgtythrrthrrthrr 120  
DB 993 AAGGACATGCAAGATGGCGGTAAAGCCCTCATGTGTGCTGGGTGGGCAAGAGCATTTA 1052

QY 121 Alaglygluglymetlellethrvalaglyylleaspserrhsthrhispheleuser 140  
DB 1053 GCAGGGAGAGGTATGATTATTAACCGCTGGGGGAGTTCATTCACACCCACTTCTCTCT 1112

QY 141 Proglnglnpheprothrvalaleuvalasnqlyvalthrthrmecphegiylgilythr 160  
DB 1113 CCACAAACATTCCTTACCGCTCTAGCCAAATGGCTTCAACCAATGTTTGGCGGTGGCA 1112

QY 161 Glyprovalasplgilythrasmalthrthrrtlethrprogllylvtrpasmleuhsarg 180  
DB 1173 GGTCCGGTAGATGGCAGAAATGCAATCACTCCCTGGGCAATGGAATTCACCGC 1232

QY 181 Metleuarglalaaglugluytrsermetasnsvalgylpheleuglylysglyasnsr 200  
DB 1233 ATGTTGGCGCGCTGAAAGATATTTCTATGATGAGCTTTTGGCGCAAGCGCAATAGC 1292

QY 201 Serseplysuglnleuvalgluglnvalgluaglyalaglylypheleuhs 220  
DB 1293 TCCAGTAAAAACAACCTGTAGAACAAATGAAAGCGGCGCGATCGGCTTAAATTCAT 1352

QY 221 Gluasptpdllythrthrrproseralalleasphiscytleuservalalaaspluytr 240  
DB 1353 GAAGACTGGGGCACACACTCCAAATGCAATGCATGCTTGAGCGTACAGATGAATAC 1412

QY 241 Aspyalginvalcysillehsthrasprthrvalasnqlyllytrvalaspsprthr 260  
DB 1413 GATGTGCAAGTTTGTATCCACCGATACCGTCAATAGGAGGTATGATGATGACACC 1472

QY 261 Leuasnalamecasnqlyargalalleaslatyrhisiilegluglyalagly 280  
DB 1473 CTGAATCGATGAACGGGGCGCGCATCATGCTCAACCAATTGAGGGGCGGGGAGGA 1532

QY 281 Hiserproaspyalillethrmetalaglyleuasnilleuaproserserthrthr 300  
DB 1533 CACTCACTGATGTATCACCATGCGAGCGAGCTCAATTTCTACCCCTCTCCACCAACC 1592

QY 301 Prothrilleprotyrrthrrtleasnthrrvalaglunhsleuaspmetleumetthr 320  
DB 1593 CCCACTATCCCTATACCATTTAATACGTTGCAACACCTTAGACATGCTCATACCTGC 1652

QY 321 Hisleuaspllyasglilearglunasplenglnpserglnserargileargpro 340  
DB 1653 CACCACTAGATTAACCATCCGCGAGATTTTACAAATTTTCCAAACCGGATCGGCC 1712

QY 341 Glyserllealagluaspyalileuhsasplmetglvalilleamethrrser 360  
DB 1713 GGCTCTATCGCGCTGAAGATGTGCTCCATGATTTGGCGGTGATCGGATGACAAAGCTCG 1772

QY 361 Aspserrglunlametgylarglvalgluvalilleproargthrtrpdlinthrvalasp 380  
DB 1773 GATTCGAGAGATGGGGCGCGCTGGCGAAGTGTCTCGAATTTGGCAACTGGCGAT 1832

QY 381 Lysasnlysblysgluphegiyllyleupproglunaspilylysaspaasnsrphsarg 400  
DB 1833 AAGAAATAAAAAGATTGGTAAGCTTCGGAAGATGTGCAAGTAAACCAACTTCGCT 1892

QY 401 Ilelysaargtyrileserlytyrrthrrtleasnsproalleuthrrhisiaglyvalserglu 420  
DB 1893 ATCAAAACGTACATCTCAATATCACATTAACCCCTCAACCTGAGGTAAAGCAG 1952

QY 421 Tyrillegyservaglugluglylvlellealasplleuvalalletpasnsproalle 440  
DB 1953 TATATCGGCTCTGGGAAGAGGGGAAATGCGATTTGGGTGGGTGAATCTGCTTT 2012

QY 441 Phegllyalilysprolysillevalllelysglylymetvalyvalpasesrglumetgily 460  
DB 2013 TTGTGTGAAACCTTAAGATTGATCAAAAGCGGTATGGGTGCTCTCTGAAATGGGC 2072

QY 461 Aspserasnalaservallprothrproglinprovalytrtyrarglumetphegiylhis 480  
DB 2073 GACTCCAAACCGGTCCGCTTACACCTCAACCGGTATTAATACCGCAAAAGTTTGGCAT 2132

QY 481 Hisgllysalalyspheasprthrserillethrphvalaserlysvlalatyrgluasn 500

Db	2133	CACGGCAAGCGCAAAATTGACACGACATCACTTTTGTTCGAAAGTCGCTATGAAAT	2192
Qy	501	GTYVAllySGlulysleuGlyleuGluValleuProValylsAsnCysatgaen	520
Db	2193	GGCGTGAAGAAAGAAACTAGAGCTTAGAGCCGCAAGTCTCCTCGTGAAGAAACGCCCAAC	2252
Qy	521	lIethrlyLeuAspPheAspPheAsnAspLysThrAlaIylelIethrValAspProLys	540
Db	2253	ATCATTAGAAAGAACTTCAAAATTCACAAACAGCGCGCATTCCTGCTGCTTA	2312
Qy	541	ThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSergluValPro	560
Db	2313	ACCTTCGAGAGCTTTGTAGATGCGAAACTCTGACCTTAACCCGCTCTGAGTGCCT	2372
Qy	561	leuAlaGlnArgTyrThrPhePhe	568
Db	2373	CTAGCCAGCGCTACCTTCTTC	2396
RESULT 8			
AX356692			
LOCUS	AX356692	2407 bp	DNA linear PAT 13-FEB-2002
DEFINITION	Sequence 10 from Patent EPI176192.		
ACCESSION	AX356692		
VERSION	AX356692.1	GI:18674029	
KEYWORDS			
SOURCE			
ORGANISM	Helicobacter felis		
	Helicobacter felis		
	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;		
	Helicobacteraceae; Helicobacter.		
REFERENCE			
AUTHORS	1 Kusters, J.G. and Catcofi, G.		
TITLE	Helicobacter felis vaccine		
JOURNAL	Patent: EP 1176192-A 10 30-JAN-2002;		
	Akzo Nobel N.V. (NL)		
FEATURES			
source	Location/Qualifiers		
	1..2407		
	/organism="Helicobacter felis"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:214"		
	2..682		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAD23261.1"		
	/db_xref="GI:18674030"		
	/translation="MKLTPKQEQKFLYYAGAVAKRKAEGLKLNPEAIAYISAHIM		
	DEARRGKTVAMLEBECMHFLKDEYMPGVGMVDPDVGVEATFPDGTKLVTNWP		
	DBHFNAGEVFGCDKDIELNAGEVTELENGEPKSLGSHFHFEBNKALEKEDRE		
	KAYGRKLDIPSGNTLRIGAGQTRKVLPIPLGSKKVIYGNGLVNNIADBRHKALEK		
	AKSHGPIK"		
	693..2399		
	/note="unnamed protein product"		
	/codon_start=1		
	/transl_table=11		
	/protein_id="CAD23262.1"		
	/db_xref="GI:18674031"		
	/translation="MKMKQEQYVNTYPTTGDKVRKLRGDDVLAWEHNDYTYGEELK		
	GAGKTRBGMGQNSPDENTLIDLVTINAMIDITGTYKADIGKNGKIHGIGAKGKD		
	MODGVPHNVAVGTEALAGKEEMIITAGGIDSHTHFLSQEQPTALANGVTMFGGT		
	GEVDGNTATITPGKWLHRLMLAAEYGMNVGFLGKSSSKQIYVEQIEAGALGFK		
	LIEDAGTTPSAIDHCLSADEYDVQCIHTDVTNBEYVDPTLNANGALIAHYHIEG		
	AGGSGPDVITMAGELNILPSSTPTPIPYTINVAEHLMLMTCHHLDRIRIBDIFS		
	QSRIRGSIADVDLHIDIVIAMTSSDSQAMGRAGEVIRRTQJADKXKEGKGPED		
	SADNDNFRIRKYSKITINPALTHGVSIVGSEBKLADLVVWNPAPFGVGVKVIK		
	GMNVFSEMGDSNASVPTPQPYVIREMFGHGGAKADITSITVSKVAYGKXKGL		
	ERKVLVKNCRNITKKDFKFNKTAHITVDPTFEVFDGKLCSTKPEAEVLAORYT		
	PP"		
CDS			
ORIGIN			
Alignment Scores:			
Pred. No.:	3..31e-187	Length:	2407
Score:	2953.00	Matches:	558

		Percent Similarity: 99.12%	Conservative: 5
		Best Local Similarity: 98.24%	Mismatches: 5
		Query Match: 98.47%	Gaps: 0
		DB: 6	Gaps: 0
	US-09-904-994b-3 (1-568) x AX356692 (1-2407)		
Qy	1	MethyMetlyLeuGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal	20
Db	693	ATGAAATGCAAAAGAAACAGAGATGATGTAATACCTTACGACCTTACACAGCGCAAGAG	752
Qy	21	ArgLeuGlyAspThrAspLeuTrpAlaGluValGluHISAspTyrThrTyrGlyGlu	40
Db	753	CGCTTAGAGATACCGATCTTGGGACAGAAATGACTATACCACTTATGCGCA	812
Qy	41	GluLeuLysPheGlyValAGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro	60
Db	813	GAGCTCAAAATTTGGCGCGGTAAACTATCCGTGAGGCGATGGTCAGAGCAATAGTCA	872
Qy	61	AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle	80
Db	873	GATGAAACACCCCTTAGATTAGTCATCACACGCGATGATTATGACTACACCGGGATT	932
Qy	81	TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn	100
Db	933	TACAAAGCCGACATTGGCATTTAAATGCAAAATCCATGCGATTGGCAAGCAGAAAC	992
Qy	101	LysAspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu	120
Db	993	AAGGACATGACAGATGGCGTAAAGCCCTCATAGTGGTGGTGGCGCACAGAACATT	1052
Qy	121	AlaGlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSer	140
Db	1053	GCGGGGAAAGATATGATTATTCCTGGGCGGATCGATTACACACCCACTTCCTCT	1112
Qy	141	ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr	160
Db	1113	CCACAACAATTCCTTACCGCTTACGCAATGCGGTAAACAACATGTTGGCGTGGCACA	1172
Qy	161	GlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysThrAsnLeuHisArg	180
Db	1173	GGTCGGTAGATGAGGACGATGCGACTACCATCTCGGGGAAATGAACTTGACCGCC	1232
Qy	181	MetLeuArgAlaAlaGluGlyLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSer	200
Db	1233	ATGTTCCGCGCGACGCTGAAGATTCATGATGTCGCTTTTGGCGAAAGCAATAC	1292
Qy	201	SerSerLysLysGlnLeuValGluGluValGluAlaGlyAlaIleGlyPheLysLeuHis	220
Db	1293	TCCAGTAAAAAACAACCTTGTAACAATAAGAAAGCGGCGCATCGGCTTTAAATTGCAT	1352
Qy	221	GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyr	240
Db	1353	GAAAGCTGGGCGACAACCTCCAGTCAACATGATCACTGCTTAGCGTGAAGATGATAC	1412
Qy	241	AspValAlaGlnValCysIleHisThrAspThrValaAsnGluAlaGlyTyrValAspAspThr	260
Db	1413	GATGCGCAAGTTTGTATTCACACCGATACCGATGAGGCGATTATGTATGATGACACC	1472
Qy	261	LeuAsnAlaMetAsnGlyValArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGly	280
Db	1473	CTGAATGCGATGACGAGGCGCCATCATGCTACCATGAGGAGCGGCGGAGGAGA	1532
Qy	281	HisSerProAspValIleThrMetAlaGlyGluAsnAlaIleLeuProSerSerThrThr	300
Db	1533	CACCTACCGATGATTAATCAACATGACGCGCTCAATATTCCTCCCTCCCAACACC	1592
Qy	301	ProThrIleProTyrThrIleAsnThrValaGluHisLeuAspMetLeuMetThrCys	320
Db	1593	CCCACTATCCCTATACCATTTAATACGTTGAGAAACCTTAGACATGCTCATGACTGC	1652
Qy	321	HisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro	340

	C12N15/09, A61K38/00, A61K39/106, A61K39/118, A61K39/12, A61K39/175, PC
A61K39/23,	
PC	A61K39/235, A61K39/39, A61K39/395, A61P1/04, A61P31/04, C12N1/15,
PC	C12N1/19,
PC	C12N1/21, C12N5/10, C12N9/80, C12Q1/58, G01N33/15, G01N33/50, G01N33/53,
G01N33/53,	
G01N33/53, G01N33/566, G01N33/569//C12N9/80, C12R1/01, C12Q1/68, PC	
C12R1/01,	
PC	C12N15/00, C12N5/00, A61K37/02
CC	Helicobacter felis vaccine
FH	Key Location/Qualifiers
FT	CDS (3) . . (683)
FT	CDS (694) . . (2181) .
FEATURES	
source	Location/Qualifiers
	1. 2183
	/organism="Helicobacter felis"
	/mol_type="genomic DNA"
	/db_xref="taxon:214"
ORIGIN	
Alignment Scores:	
Pred. No.:	5.6e-163 Length: 2183
Score:	2588.00 Matches: 489
Percent Similarity:	99.59% Conservative: 1
Best Local Similarity:	99.39% Mismatches: 2
Query Match:	86.30% Indels: 0
DB:	6 Gaps: 0
US-09-904-994B-3 (1-568) x BD185304 (1-2183)	
OY	1 MetLysMeTtLYSLySGInLUrYrVAlAsnThrTYGlyProThrLySGlYAspLYsVal 20
Dd	694 ATGAATAAGAAAAACAAGATATGTAAACACTAGCGACCACCAAGCGATTAAAGTG 753
OY	21 ArGLeuGIYAAsPTnAsPLEutTPAlaGlUvAlGUhLSAsPTYrThrTYrGLyGLU 40
Dd	754 CGCTTAGAGATACCGCATCTTTGGCAGAGAAGTAACAATGACTATTCACCTTAGCGAA 813
OY	41 GLIueULySPheGLYAlaGLYLySrThrlleaRGluGLIMetGLYGINSerAsnsErPro 60
Dd	814 GAGCTCAAATTTGGCGCGGGTAAACTATTCGTAGAGGTATGGGTCAAGCAATACCCA 873
OY	61 AsPGlUAsnThrLeuAsPLeuVallleThraSnAlaMetlleleAsPTYrThrglylie 80
Dd	874 GATGAAAACAACCTTAGATTAGTATGCACCAACGCGATGATTATCGACTAACCGGAAAT 933
OY	81 TyrtLysAlaAsPIllEGlytlleLysasnGLYlysllleHlsGLYtllegLYLysAlaGLYaen 100
Dd	934 TATAAAGCGCAATTGGTATTAATAAATGCAAAATCCATGATTTGGCAAGCGGGGAAC 993
OY	101 LYSAsPMetGINAsPGlYValSeErProHisMEtValAlGLYValGLYThrGLUALaleu 120
Dd	994 AAAAGCAICGAAGAGGGCGTAAGCCCTCATATNGTCGTGGGTGGGCACAGAAGCACTA 105
OY	121 AlAglyGLUGlyMetllelleThralagLYtlleAsPSerHisThriSphelEusEr 140
Dd	1054 GCAGGGGAAGGATATGATTAATTAACCCCTGGGGGATCGATTCACACACCCACTTCCCTCT 111
OY	141 ProGLINGlnPheProThralAleuAlaasnGLYValllThrThMeTPhEGlyGLyGLYThr 160
Dd	1114 CCCCAACAATTCCTTACCGCTTACGCCAATGTGTTCACACCAATGTTTGGAGGTGCACA 117
OY	161 GlYProVALAsPGlYThraSnAlaThrThrlleThrProGLYLySTTPAsnLuhIsarG 180
Dd	1174 GGTCGGGTAGATGGCACCAATGCGCACCACTACCTCGGGCAAAATGGAACTTGACCCC 123
OY	181 MetLeuAsgAlaaGLUGlyTSerSerMetAsnVALyPheLeuGLYLySGlYAsnsEr 200
Dd	1234 ATGTTGCCCGCGAGTAAGATATTCTTAAGATGTAGGCTTTTGGCAAGGCAATAGT 129
OY	201 SererLYSLySGInLeuValGUcInlVaAlGUlaGLYAlalleGLYPheLYSLeuHis 220



Db	1294	TTTAGCAAAAAACA	CTTACAGACAGTAGAAGCGGCGCATTTGGCTTTAAATTGCAT	1353
Qy	221	GIUAAPRTGTYL	ThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyr	240
Db	1354	GAAAGACTGGGGCA	CAACCAAGTGGCATGCATCTTGAGCGCTGGCAGATGAAATAC	1413
Qy	241	AAAPVALGlnVal	CysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThr	260
Db	1414	GATGTGCAAGTTT	GTATTCACACCGCATACCGCTCATAGAGCGAGTTATGTGATGACACC	1473
Qy	261	LeuAsnAlaMetAsn	GlyArgAlaIleHisAlaTyrHisIleGluGlyValaGlyGly	280
Db	1474	CTAAATGCAATGA	TACAGCGGCGCGCATTCATTCCTTACCATTTGAGGAGCGGCGGAGA	1533
Qy	281	HisSerProAspVal	IleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThr	300
Db	1534	CACCTACCTGATG	TATTCACATGGACGGCGAGCTCAATATTTCTACCTCTCCACACCC	1593
Qy	301	ProThrIleProTyr	ThrIleAsnThrValAlaGluHisIleAspMetLeuMetThrCys	320
Db	1594	CCCATATTCCTT	ATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1653
Qy	321	HisHisIleuAsp	LeuArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro	340
Db	1654	CACCACTGATTA	AGCGGATCCGCGAGATTACATTTTCTCAAGCCGTATCCGCCCC	1713
Qy	341	GlySerIleAlaAla	GluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer	360
Db	1714	GGATCTATTGCG	CGCTGAGATGTGCTCCATGATATGGCGTATCCGATGATCTACCTCC	1773
Qy	361	AspSerGlnAlaMet	GlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAsp	380
Db	1774	GATTCGCAAGCA	TATGGGGCGCTGGGGAAGTATTCCTAGAACTTGGCAACTGCAAGT	1833
Qy	381	LysAsnLysLysGlu	PheGlyLysLeuProGluAspGlyLysAspAsnAspAsnAsp	400
Db	1834	AAGAATTAATAA	AAAGATTTGTGTAAGCTTCTGAAGATGTGCAGATACACCACTCCGC	1893
Qy	401	IleLysArgTyrIle	SerIleThrIleAsnProAlaLeuThrHisGlyValSerGlu	420
Db	1894	ATCAAAACGGTAT	ATATCTCCAAATACACCATTAATCCCGCTTTGACCCATGGCGGAGCGAG	1953
Qy	421	TyrIleGlySerVal	GluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhe	440
Db	1954	TATATCGGCTCT	GTGTAAGAGGCAAGATCGCGAATTGGTGTGGAATCTCGCTTT	2013
Qy	441	PheGlyValLysPro	LysIleValIleLysGlyMetValValPheSerGluMetGly	460
Db	2014	TTTGGCGTAAAC	CTTAAGATTGATTAAAGGTGCGATGTGGTCTTCTGAAATGGGC	2073
Qy	461	AspSerAsnAlaSer	ValProThrProGlnProValTyrTyrArgLysMetPheGlyHis	480
Db	2074	GATTTCAAGCGC	CTCGCCACGCTCAGCGCTTATTAACCGCAAAATGTTTGGGCAC	2133
Qy	481	HisGlyLysAlaLys	PheAspThrSerIleThrPhe	492
Db	2134	CACGGCAAGCGA	AAATTGACACACACATCACTTTT	2169
RESULT 10				
AX356689				
LOCUS	AX356689	2183 bp	DNA	linear
DEFINITION	Sequence 7 from Patent EP1176192.			
ACCESSION	AX356689			
VERSION	AX356689.1	GI:18674026		
KEYWORDS				
SOURCE	Helicobacter felis			
ORGANISM	Helicobacter felis			
REFERENCE	1			
AUTHORS	Kusters, J.G. and Cantolli, G.			
TITLE	Helicobacter felis vaccine			

JOURNAL	Patent: EP 1176192-A 7 30-JAN-2002;
FEATURES	Akzo Nobel N.V. (NL)
source	Location/Qualifiers
	1..2183
	/organism="Helicobacter felis"
	/mol_type="unassigned DNA"
	/db_xref="taxon:214"
CDS	3..683
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23259.1"
	/db_xref="GI:18674027"
	/translation="MKLTPKOEKFLVYAGVARKRAEGLKLNQPALAYISAHIM
	DEARGRKTVAEIMEECMPLKDEVMGVGMVDLGVFAFPFGTGLVTVMNLEP
	DEHFRAGYKRGCCDDIETLNQKSTVETLEVTNEGKSLAHSGHFHFFETNKLKRE
	KAYGRKDIPEGNTLRIGAGQTRKQLPLGGSKVITGMNGLVNNIADBRHKKALDK
	AKSHGPIK"
	694..>2181
	/note="unnamed protein product"
	/codon_start=1
	/transl_table=11
	/protein_id="CAD23260.1"
	/db_xref="GI:18674028"
	/translation="MKMKOEYVNTYPTGDKVRLGDTDMAEVEDHYTYEELKF
	GAGKTIREGMGOSNPDENTLDLVTNNMIIIDYTGIRADIGIKNGKIHGIGAKND
	MODGVSPHVVGVTEALAGSMITTAGISDHTFPLSPQEPPLALANGVTMPGGT
	GPVQSTNAFTITTRGKWNLAHRLRAEEYSNVVGPIGKNSSSKKQLVQVEVAGIFK
	LHEDGTTPSAIDHCLSVADIEDVQVCIHTDTVNEAGVYVDITLNMNGRAIHAHIBG
	AGGGSPDVTIMAGELNLPSSSTPTIYTNVTAHEHIDMLMTCNHLKREDLQFS
	OSRIRPDSIAEDVDLHDIGVIAMTSSDSQAMRAGEVLPRTWQTDKNKKEPKLVIK
	GADNDFRIKRYISKYTINPALVTHSVSYSGVEBGIADLVVNNPAPFGVGPXKIVIK
	GMVVFSEMGDSNBSVPTPPQVYVEMEGHHKAKFDSIFRVS"

## CDS

Alignment Scores:					
Pred. No.:	5..6e-163				2183
Score:	2588..00				489
Percent Similarity:	99.59%				1
Beet Local Similarity:	99.39%				2
Query Match:	86.30%				0
DB:	6	Gaps:			0

## ORIGIN

US-09-904-994b-3 (1-568) x AX356689 (1-2183)					
Qy	1	MetLysMetLysLysGluGlnTyrValAsnThrTyrGlyProThrLysGlyAspLysVal			20
Db	694	ATGAATAATGAATAAACAAGATGATGTAAACACTACGACCCACACAGCGCATTAAGTG			753
Qy	21	ArgLeuGlyAspThrAspLeuTrrAlaGluValaGluHisAspTyrThrTyrGlyGlu			40
Db	754	CCCTTAGAGAGATACCGATTTTGGGCAGAGTAGAACATGATTAACATTAATGCGGA			813
Qy	41	GluLeuLysPheGlyAlaGlyAspThrIleArgGluGlyMetGlyGlnSerAsnSerPro			60
Db	814	GAGCTCAATTTGGCGCGGTAAACTATCCCTGAGGATATGGTCAGAGCAATAGCCCA			873
Qy	61	AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle			80
Db	874	GATGAAACACCTTGATTTAGTATGATCAACCAAGGAGATTAATGACATCACCCGGATT			933
Qy	81	TyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn			100
Db	934	TATTAAGCGACATTTGTTTAAATGCAAAATTCATGATTTGGCAAGCGGGGAAC			993
Qy	101	LysAspMetGlnAspGlyValSerProHisMetValaGlyValaGlyThrGluAlaLeu			120
Db	994	AAAGACATGACAGATGCGGTAGCCCTCATATGTGTGCGGTGGGCGACAGACACATTA			1053
Qy	121	AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer			140
Db	1054	GCAGGGGAAGATATTAATTAACCGCTGGGGGAGATCGATTCGACACACCATTCCTCTCT			1113



141 ProGlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThr 160  
 1114 CCCCAACATTCCTTACCCTCTTACCAATGGTCTTACACCAATGTTGGAGTGGCA 1173  
 161 GlyProValAspGlyThrAsnAlaThrThrLeThrProGlyLysTrpAsnLeuHisArg 180  
 1174 GGTCCGGTAGATGGCAGAAATGGCACCACCATCTCCGGGCAATGGAATCTTCAACGC 1233  
 181 MetLeuArgAlaAlaGluGlyTrpSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 200  
 1234 ATGTGGGCGCACGCTGAAGATATTCATGATATAGACTTTTGGCGCAAGGCAATAGT 1293  
 201 SerSerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLeuLeuHis 220  
 1294 TCTTACCAAAAACACTTGTAGAACATGACACGGGCGGCTTAAATTTGAT 1353  
 221 GluAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyLys 240  
 1354 GAAGACTGGGGCACAACCAAGTGCATGCATCTGACGCTTGAACGTCGAGATGAAATAC 1413  
 241 AspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTrpValAspAspThr 260  
 1414 GATGTGCAAGTTGTATCCACCGATACCGTCAATGAGCGAGGTATGTGATGACACC 1473  
 261 LeuAsnAlaMetAsnGlyArgAlaIleHisAlaThrHisIleGluGlyAlaGlyGly 280  
 1474 CTAAATGCAATGAACGGGCGGCGCATGCTTACCAACATGAGGAGGCGGCGGAGGA 1533  
 281 HisSerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThr 300  
 1534 CACTCAGCTGATGTTATCACCATGGCAGGAGGTCAATATTCACCTCTCCACCAACC 1593  
 301 ProThrIleProGlyThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCys 320  
 1594 CCCACATTCCTCATACCATTAATACGTTGCAACACATTAAGCATGCTCATGACCTGC 1653  
 321 HisHisLeuAspLysAspArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgPro 340  
 1654 CACCACTACATTAAGGCATCCGCGAGATTTACAAATTTCTCAAGCCGATATCCGCCCC 1713  
 341 GlySerIleAlaIleGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSer 360  
 1714 GATCTATTCCTCCGCTGAGATGCTCCATGATATGCGCGATGCGGATGATGATGCTCC 1773  
 361 AspSerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAsp 380  
 1774 GATTCGCAACCAATGGGCGCGCTGGGGAAGTGAATTCCTGAACCTTGCAAACTGCAGAT 1833  
 381 LysAsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArg 400  
 1834 AAGAAATTAATAAAGATTTGGTAACTTCTGAAAGATGTCAGATTAACGACAACTTCGCG 1893  
 401 IleLysArgTrpIleSerLysTrpThrIleAsnProAlaLeuThrHisGlyValSerGlu 420  
 1894 ATCAACGCTATATCTCAATACACCATTAATCCGCTTGAACCATGGCGTGAAGAG 1953  
 421 TyrIleGlySerValGluGluGlyLysIleAlaAspLeuValValITPAsnProAlaPhe 440  
 1954 TATATCGGCTCTCTGGAAGAGGCGAGATGCCCACTTGTTGGTGGAAATCCCTGCTTT 2013  
 441 PheGlyValLysProLysIleValIleLysGlyGlyMetValValIPheSerGluMetGly 460  
 2014 TTTGGCGTGAACCTTAAGATTTGATTAAGTGGCAATGGTGGCTTCTCTGAAATGGGC 2073  
 461 AspSerAsnAlaSerValProThrProGlnProValTyrTrpArgGluMetPheGlyHis 480  
 2074 GATTCATACGCGCTCGGCGCCAGCGCTCAAGCGGTTATTAACGCGAAATGTTTGGCAC 2133  
 481 HisGlyLysValLysPheAspThrSerIleThrPhe 492  
 2134 CACGGCAAGCGAAATTTGACACGACGATCACTTTT 2169

RESULT 11

AF330621  
 LOCUS 8406 bp DNA linear BCT 13-MAY-2003  
 DEFINITION Helicobacter bizozeronii tRNA ribosyltransferase gene, partial cds; GidB, complete cds; urease gene cluster, complete sequence and unknown gene.  
 ACCESSION AF330621  
 VERSION AF330621.1 GI:27462193  
 KEYWORDS Helicobacter bizozeronii  
 SOURCE Helicobacter bizozeronii  
 ORGANISM Helicobacter bizozeronii; Epsilonproteobacteria; Campylobacterales; Helicobacteraceae; Helicobacter.  
 REFERENCE 1 (bases 1 to 8406)  
 Zhu, J., Teng, C.H., Chang, C.F., Chang, C.D., Simpson, K.W., Wei, C., McDonough, P., McDonough, S., and Chang, Y.F.  
 Cloning and characterization of a Helicobacter bizozeronii urease gene cluster  
 JOURNAL DNA Seq. 13 (6), 321-331 (2002)  
 MEDLINE 22540228  
 PUBMED 12652903  
 REFERENCE 2 (bases 1 to 8406)  
 Zhu, J., and Chang, Y.F.  
 Direct Submission  
 TITLE Submitted (21-DEC-2000) Population Medicine and Diagnostic Science, College of Veterinary Medicine, Cornell University, Tower Road, Ithaca, NY 14853, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source  
 1. 8406  
 /organism="Helicobacter bizozeronii"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:56877"  
 <1..929  
 /codon\_start=3  
 /transl\_table=11  
 /product="tRNA ribosyltransferase"  
 /protein\_id="AAO15370.1"  
 /db\_xref="GI:27462194"  
 /translation="LMVYERASAKLTHSTFKEIPDPPRHVALVNDTKYIKARLPH TAQRREPEILHRTSPRTACLAOMRGKVRGVLGDBGYSCEVLEFVLDGLVLSFK HGGELMEAGVLEMLBLGHVPLPYLKRDEQDQDYOSVAKYFGALAAAPASLH FQSADHLLKNKFHVFITLVGAGTFLSYOSDIDIEHPMAFCRSLPQAGLTLEDA PYLCVGTALRATHEHYKRGFGKPELFLHLGNPVYHAQALLTNFHLPESTLIMLVA SWGLDKCELYKATIEBHHRFYSYDGMIL"  
 926..1465  
 /codon\_start=1  
 /transl\_table=11  
 /product="GidB"  
 /protein\_id="AAO15371.1"  
 /db\_xref="GI:27462195"  
 /translation="MRKADLPTHLLEMGSVANLSGAKKHODIEHNIDSLQVDFI APPKSLDLSGAGFPALPISLACNARFLLPNAKRAVFLRHVKALNLNLELOR MRIEHSPOSIVLDITSRALMAONLIALSAPLEQGHFLFYKSHLRTETIACDH ECHVYGRVYFYQQRDFA"  
 1670..2053  
 /codon\_start=1  
 /transl\_table=11  
 /product="unknown"  
 /protein\_id="AAO15372.1"  
 /db\_xref="GI:27462196"  
 /translation="MLVFGHPOICPMFKCVADIOAIVPSHTIPFQPCHTDAFL SKKCSQGVYAVYVQNVNPLKCALMAFKPTIYLAENQLACGSIANDYLDPAILKYI EHAQLEBIVKILRDGAVRSFLML"  
 2341..3045  
 /codon\_start=1  
 /transl\_table=11  
 /product="urease A"  
 /protein\_id="AAO15373.1"  
 /db\_xref="GI:27462197"  
 /translation="MKLTPELDKMLHYAGELAKKRKANGKVCNLTAAVALISAHV BEARPAKGSVADLMQGRLLKADVMPGVANIHVGIEANPDETKVTITPVED GGHKLAKGEVILNEDITLNAKQATTLLENKGDHAPVGVSHFHFVVKLLEPRE KAYGRDLIASGTVARFEPQEKTEVLIQIGGNRIYFNLSLVDRQADTDGKKLAKR AKHGRGVNCGDKK"

CDS  
3061..4770  
/codon\_start=1  
/transl\_table=11  
/product="urease B"  
/protein\_id="AA015374.1"  
/db\_xref="GI:27462198"  
/translacion="MKISRKEYVSMYGPFTTDKVRLEGDTDLLEVEHDCCTTYGEEIK  
FGGKTRIDGMQNTSPSSHELDLVNMLIVDYGIYKADIGIKNGIKHIGKAGNK  
DMODVCNNICVGPATREALAAGLIVTAGGIDITHHIFTSPQIPRAFGSGITMMGG  
TGPADNATTTTPGRNMLKTMRLRASEYANMLGYLKGKNSVEPSLVNDQLEAGATG  
KIHEDWGSTPAIYHCLNVAADKYDVQVALIHDTLNEAGCVEDTLQALIGRTITHTTE  
GAGGHPADVIKMSGEFNILPASTNPILPTVNTAENHDMVCHLIDKNTKEVOP  
ADSRIRPOTIAEDKLDMDGIFSTSSDQAMRGVEVITRTMQADKKKKEGFLPS  
EKGDNDFRIKRYISKYITNPALNPAISYSGSVSGVADLVMSPAFGKPKMTI  
KGFIALSGMDANASISIPPOPVYRAGMGHGHKAKPDNITPVSQVAVDNGIKKEIG  
LORVULPVKNCNITKKDLKPDVTAHIEVNPETYKVKVKGSEVTSKAADKISLQVLY  
NLIF"

CDS  
4965..5552  
/codon\_start=1  
/transl\_table=11  
/product="urease 1"  
/protein\_id="AA015375.1"  
/db\_xref="GI:27462199"  
/translacion="MGLVLYVAIVLISNGVCGLANVDARSKAVMNVFVGSLIICN  
VIAIYVSTENPTPTVGAEDVAQVSOQLINFPYGPATGLFGETYMAAINNIYNLDMK  
PYGVCLEFVITNTPAIIISHYSDALDHRILGITEGDMWAFITLWAMGLMTWMEIC  
ALGSLGKVPPLAIIIEGVITAMIPAMLIIFIQMS"

CDS  
5563..6093  
/codon\_start=1  
/transl\_table=11  
/product="urease B"  
/protein\_id="AA015376.1"  
/db\_xref="GI:27462200"  
/translacion="MLASVILGNLKEGSSKELPDLEMPDQKMRGPTSSOKAEI  
VLAKPRPMGLCDGDIPEBDTSLAINITPTPLHYADSTAOVARKCTEYGRHA  
SLYGDSPLSFETPERPLQVLFDKALKRYELSKLDASQRI SVSAPHADLQSGSA  
PLKFKSALDLQIVIRK"

CDS  
6098..6775  
/codon\_start=1  
/transl\_table=11  
/product="urease F"  
/protein\_id="AA015377.1"  
/db\_xref="GI:27462201"  
/translacion="MQDFLLQVNDAMPFGSYTHSFGLETYIOHKEVESKESALDY  
MRAVSTOFLYTELLAKLYAGVAHQEDLESILMEMBOICLATPPELEBSANOKGNR  
FLKTIQVLDPLSAPFKDVMOKSTPTTHASAVGCACIGLSELSIKHYLYAOSXNM  
VINCXYTIPLAQNDQRIILALQETFTTLVTLLEGUDASYLCAASI QNDIKAMQHEHL  
YSRLYMS"

CDS  
6797..7396  
/codon\_start=1  
/transl\_table=11  
/product="urease G"  
/protein\_id="AA015378.1"  
/db\_xref="GI:27462202"  
/translacion="WVKIGVCGPVSGKLTALIEALTTRANSQOYSIGVITNDIYTKEDA  
EFLCANSWPRERIIGVERGCGPHTAIRDASNMLEAVELEIAKPEIDLEIMFI EESGD  
NLSITFENPELADFTIFVIDVAGDKIPKRGSGITRSDLLINKIDLAIPYVADLGYM  
DRDSKMGDKRFLFTNIRSKGDLHVIGWIKKVALLEB"

CDS  
7399..8232  
/codon\_start=1  
/transl\_table=11  
/product="urease H"  
/protein\_id="AA015379.1"  
/db\_xref="GI:27462203"  
/translacion="MAVGHPTSYAQAPRLYLRTKIQNGKCYAIADNYSPPKLMPF  
VERGMABETILIAVSPGMLGDAQEI QMDIGENCQLELSFQSEKIQDTEDEYATANT  
QHIIHNLADSPLEPIIPANAHFKNSHIVLHNAQOLVSEIITAGRIANGEFAP  
HIIHSTLKISTQEGESRSVFLDNTILBASMDLKNPCMFQVYTHYLNIFITRLILKL  
EDLEHLESTQINAGVSLPTRLPGYHSLCLKALANGSEPLALRKSISNLVVERME  
G"

ORIGIN  
Alignment Scores:

Pred. No.: 7,766-145 Length: 8406  
Score: 2326.00 Matches: 420  
Percent Similarity: 87.30% Conservative: 75  
Best Local Similarity: 74.07% Mismatches: 72  
Query Match: 77.56% Gaps: 0  
DB: 1 Indels: 0

US-09-904-994b-3 (1-568) x AF330621 (1-8406)

QY 2 LysMetLysbysGlnGluTyrValaLenThrTyrGlyProThrLysGlyAspLysValArg 21  
DB 3067 AAAAATCTCTCGAAAGAAATATGTTCTATGTATGAGACCACTACGCGCATTAAGTGA 3126

QY 22 LeuGlyAspThrAspLeuTrpAlaGluValaGluHisAspTyrThrThrTyrGlyGlu 41  
DB 3127 TTGGGCGATCCGACCTGATCTTAAGAAGTCAACATGATGACCACTTATGCGAAGA 3186

QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61  
DB 3187 ATTAAGTTTGGTGGCGGTAAACCATTCGCGATGGGATGGCACAAACCAAGCCCGAGC 3246

QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81  
DB 3247 AGCCACGAATCGATCTGTGCTCAACCGCCGATCGTGAATTACACCGGCATTTAT 3306

QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101  
DB 3307 AAGCCGATTTGGCATTTAAATGCAAAATCCATGCGATTGGCAAGCAGCATATA 3366

QY 102 AspMetGlnAspGlyValSerProHisMetValValaGlyValaGlyThrGluAlaLeu 121  
DB 3367 GACATGCAAGATGGGCTTGGCAACATCTTGGCGCCCTGCTACTGAGGCTTTGGCC 3426

QY 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141  
DB 3427 GCTGAAGGCTGATTTGTAAGCTGTGGATGTGACCCACATCCATCTTATTTCTCCC 3486

QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyIleThrGly 161  
DB 3487 CAACAATCCCAACAGCATTTGCCAGCGGATCAACAATCATGATTTGGGGAGACAGGT 3546

QY 162 ProValaAspGlyThrAsnAlaThrThrIleThrProGlyLysTyrAsnLeuHisArgMet 181  
DB 3547 CAGCTGATGGGATCAAGCGCATCACTCCGCGGCGCTGGAACCTTTAAACCAT 3606

QY 182 LeuAlaGluAlaGluGlyTyrSerMetAsnValaGlyPheLeuGlyLysGlyLysSerSer 201  
DB 3607 CTCCGTGCTCTGAAGAAATATGCCATGAACTTGGGCTATTTGGTAAAGGAGATGTCT 3666

QY 202 SerLysLysGlnLeuValaGluGlnValaGluAlaIleGlyPheLysLeuHisGlu 221  
DB 3667 TATGAACCTCCCTGTGTGATTAACCTGAGAGTGAAGCATTTGGCTTTAAATCCAGAA 3726

QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241  
DB 3727 GACTGGGCTAGACACCTGAGCCATCACTTCTTGAATGTGGCTGACAAATACGAT 3786

QY 242 ValGlnValaCysIleHisThrAspThrValaGlnGluAlaGlyTyrValaAspAspThrLeu 261  
DB 3787 GTGCAAGTGGCTATCCACCGCATTCCTGAATGAAGGGGGTGTGTGGAAGACACTTGG 3846

QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGlyLysHis 281  
DB 3847 CAAGCATTTGCTGGGCGCATATCCACACTTCCACATGAAAGTGTGTGGCGGGCAGC 3906

QY 282 SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrThrPro 301  
DB 3907 GCTCCGAGTGTCAATTAAGATGTGGCAATTTAATCATCTCCAGCTTCTCCCAACCC 3966

QY 302 ThrIleProTyrThrIleAsnThrValaGluHisLeuAspMetLeuMetThrCysHis 321  
DB 3967 ACCATTTCTTCCACCGTGAATACAGAACCCGAACATGAGATGTGATGTGTGCCAC 4026

Qy 322 HlaeubaplysaargllleargluepgleupheseSerInserArgllleargProgly 341  
Db 4027 CACTGGATTAATAAATCAATCAAGAAAGATGTCCAGTTGTGATTTAGAGATTGCCCCCAA 4086  
Qy 342 SerIleAaIaagluapvalleuHlaaapmetGlyValleAaMetTherSerasp 361  
Db 4087 ACCATGCCCGCTGAGCAAACTCCACGATATGGGAAATTTCTATATCACACGCTTGAC 4146  
Qy 362 SerGlnAaMetGlyValleAaGlyValleProargThrTrpGlnThraAaPlys 381  
Db 4147 TCCCAACCGATGGCGCTGAGGAGGATTCATCCCGCATCTTGCAAAACGCGACAAA 4206  
Qy 382 AsnLysLysGluPheGlyLysLeuProGluaspGlyLysAspAsnAspAspPheArgIle 401  
Db 4207 AACCAAAAGAAATTTGGTCGCTTGAGGAAAAGCGCAATATGACAACTTCGCGATC 4286  
Qy 402 LysArgTrpIleSerLysTrpThraaPnProAlaLeuThraIleGlyValSerGly 421  
Db 4267 AAGGCTACATTTCCAAATACACATCAACCCGCTATGCAACGCGATTTCTGAATAT 4326  
Qy 422 IleGlySerValGluGluGlyLysIleAaAspLeuValValTrpAsnProAlaPhePhe 441  
Db 4327 GTGGGCTCTGAGAAAGTGGCAAAATGCGCATTTGGCTTGAGAGCTGCGTCTCTT 4386  
Qy 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAsp 461  
Db 4387 GGCAATTAACCAACATGATCATCAAAAGCGGATTCATCGCATTTCTCAATAGGCGCAT 4446  
Qy 462 SerAsnAaSerValProThrProGlnProValTrpTrpArgGluMetPheGlyHis 481  
Db 4447 GCCAATGCTCTATNCCCACTCCCAACCCGCTATTAACCGCAAAATGTTGGCACCAT 4506  
Qy 482 GlyLysValLysPheAspThrSerIleThraPheValSerValAlaTrpGluAsnGly 501  
Db 4507 GGTAAGCCCAATTTGACCAATATCATTTGTATCCCAAGGGCTTATGAAACAGCGC 4566  
Qy 502 ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnIle 521  
Db 4567 ATTAAAGAAAGTGGCTTGCAAAAGTGTGTTGCAATTAATAACTGCGCAACATC 4626  
Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThraLysAspProLysThr 541  
Db 4627 ACCAAAAGACCTCAATTCACATGATGTTACCGCACACATCGAAGTCAATCCGAAAC 4686  
Qy 542 PheGluValPheValLysGlyLysLeuCysThraSerLysProThrSerGlnValProLeu 561  
Db 4687 TACAAAGTTAAAGTGGTGAAGAGAGTTACTTCAAAAGCAGCGATTAATACAGCTTA 4746  
Qy 562 AlaGlnArgTrpThrPhePhe 568  
Db 4747 GCACAACTTAACAATTGTTTC 4767

RESULT 12  
AY295085 1710 bp DNA linear BCT 10-JUN-2003  
LOCUS Helicobacter pylori urease B (ureb) gene, complete cds.  
DEFINITION  
ACCESSION AY295085  
VERSION AY295085.1 GI:31580720  
KEYWORDS  
SOURCE  
ORGANISM Helicobacter pylori  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
Helicobacteraceae; Helicobacter.  
REFERENCE  
1 (bases 1 to 1710)  
TITLE Duan, G.C. and Dai, L.P.  
The ureb gene sequence of Helicobacter pylori strain isolated from  
JOURNAL Unpublished  
2 (bases 1 to 1710)  
AUTHORS Duan, G.C. and Dai, L.P.  
TITLE Direct Submission  
JOURNAL Submitted (11-MAY-2003) Epidemiology, College of Public Health,  
Daxue Road 40th, Zhengzhou, Henan 450052, P.R. China

FEATURES  
source location/Qualifiers  
1..1710  
/organism="Helicobacter pylori"  
/mol\_type="genomic DNA"  
/strain="MEL-HP27"  
/db\_xref="taxon:210"  
/country="China"  
/note="authority: Helicobacter pylori Duan G.C."  
1..1710  
/gene="ureB"  
1..1710  
/gene="ureB"  
/gene="ureB"  
/codon\_start=1  
/transl\_table=1  
/product="urease B"  
/protein\_id="AAP51176.1"  
/db\_xref="GI:31580721"  
/translation="MKKISRKEVSMVPTTGDKVRLGDTDLAEVHDYTYGEBLK  
FGGKTLRBSGMSNPSKEELDLITNALIVDTYGVKADIGIKKQKINGIGGAK  
DMQDGVKNISVGPATBALGEGITATAGIDTHIHIFISPOQIPFAPASGVTTITGG  
TGPDADGYNATITTPGRNRLKMLRAEYSMNIGFLAKGNASNDASLADQIEAGATG  
KIHBDWCTPSALNHADVDKDYVAIHITDLNBAECVEDMMAIAGTMTFHE  
GAGGCHAPDIIKYAENHILPASTNPTPTVNTAEAHMDMLMCHLIDKSIKEDVDF  
ADSRIRPQTIABEDTLHDGIFSTISDSQAMRGVEGVITRTQTDANKKEFRLKE  
EKGDNDNFRIRKYLSTKTYTNPAIAGHISEVSGVEGVKADIVLMSAPFVGKPNMTI  
KGGFIALSQGDANASIPTPQPYTRMEFPHGAKAKDANITFVSKAYIDKIKELG  
LEQVIFVKNCRNITKDMQFNDTNAHIEVNPETHVFDGKEVTSKPRKVSIAQLF  
SIF"

ORIGIN  
Alignment Scores:  
Pred. No.: 5.14e-144 Length: 1710  
Score: 2301.00 Matches: 417  
Percent Similarity: 86.77% Conservative: 75  
Best Local Similarity: 73.54% Mismatches: 75  
Query Match: 76.73% Indels: 0  
DB: 1 Gaps: 0

US-09-904-994B-3 (1-568) x AY295085 (1-1710)

Qy 2 LysMetLysLysGluGluValleThraaPnProAlaLeuThraIleGlyValSerGly 21  
Db 7 AAGATTGCAAGAAAGAAATGTTCTTATATGCGCTTACAGCGCAATTAAGTGA 66  
Qy 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTrpThrTrpGlyGluGlu 41  
Db 67 TTGGGCGATACAGCTTATCGCGTGAAGTGAACATGATCAACCATTTATGGCGAAG 126  
Qy 42 LeuLysPheGlyLysGlyLysTrpIleArgGluGlyLysGluGluGluGluGluGlu 61  
Db 127 CTTAAATTCGATGGCGGTAAACCTTGAGAGAAAGGCATGAGCCATCAACACCTTAC 186  
Qy 62 GluAsnThrLysAspLeuValleThraaPnProAlaLeuThraIleGlyValSerGly 81  
Db 187 AAGAAAGAACTGATTTATATCATCATCAACGCTTATATCGGATTTACACCGGATTTAT 246  
Qy 82 LysAlaAspIleGlyLysLysAsnGlyLysIleHisGlyLysGlyLysAlaGlyAsnLys 101  
Db 247 AAAGCGGATATTGGTATTAAGATGCAAAATCGCTGCGATTTGGCAAAAGCGCGCAAAA 306  
Qy 102 AspMetGlnAspGlyValSerProHisMetValValGlyValGlyThraGlnAlaLeuAla 121  
Db 307 GACATGCAAGATGGCGTTAAACCAATCTTAAAGCGTGGTCTGCTACTGAAGCCTTAC 366  
Qy 122 GlyGluGlyMetIleIleThraIleGlyLysLysSerHisIleThraIlePheLysSerPro 141  
Db 367 GGTGAAGGTTGATCGTAACATGCTGCGGTATGACACACACATCCACTCATCTCCGCC 426  
Qy 142 GlnGlnPheProThraIleAlaLysGlyValThraMetPheGlyGlyGlyThraGly 161  
Db 427 CAACCAATCCCTACAGCTTTTGCAAGCGGTGAACAACGATGATGTTGGTGGCGGAACCTGC 486  
Qy 162 ProValAspGlyThrAsnAlaThrThrIleThraProGlyLysTrpAsnLeuHisArgMet 181

Db CCTGCGATGGCAGCTTAACCAACCACTATCCAGCGAGAAATTTAAATGAGG 546  
Qy 182 LeuATGAlaAlaGlUglUtrSerMetAanValGlyPheLeuGlyLySGLyAanSerSer 201  
Db 547 CTCAGAGCGCGTGAAGAAATTTCTATGAATTTAGGTTTCTTAGCTAAAGTAAACGGCTTC 606  
Qy 202 SerLySlyGlnLeuValGlnGlnValGlnAlaGlyAlaIleGlyPheLeuSerLeuHISGlu 221  
Db 607 AATGATCGAGCTTACCGCATCAATGAGCGGAGCGGATGGCTTTAAATCCATGAA 666  
Qy 222 AspTrpGlyThrThrProSerAlaIleAspHISGlyLeuSerValAlaAspGlyUtrAap 241  
Db 667 GACTGGGGAACAACCTCTTCTGCAATCATCATGCGTTAGATGGTCGGAGCAAAATGCAT 726  
Qy 242 ValGlnValCysIleHisThrAspThrValAsnGlnAlaGlyTyTyValAspAspThrLeu 261  
Db 727 GTGCAAGTCGCTATCATCGGACACTTGAATGAAAGCGGTTGTGTAGAGACACTATG 786  
Qy 262 AanaIaMetAanGlyAArgAlaIleHisAlaTyHisIleGlnGlyValaGlyGlyVHIS 281  
Db 787 GGAGCCATTCGCCGAGACACTATGACACTTTCACACTGAAGCGCTGCGGAGAC 846  
Qy 282 SerProAspValIleThrMetAlaGlyGlnLeuAanIleLeuProSerSerThrThrPro 301  
Db 847 GCTCGATATCATTTAAAGTAAGTACGCGCGGAGCAACAACATTCGCGCTTCCACTAACCC 906  
Qy 302 ThrIleProTyThrIleAsnThrValAlaGlnHisLeuAspMetLeuMetCysVHIS 321  
Db 907 ACTATCCCTTTTACTGTGAATACAGAAAGCAACATGACACTGTCTTAGTGTCAC 966  
Qy 322 HisLeuAspLyAArgIleArgGlnAAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
Db 967 CACTTGGATTAAGCATTTAAAGAAATGTCACGTTGCTCATTCAGAAATCCGCTTCAA 1026  
Qy 342 SerIleAlaAlaGlnAAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAap 361  
Db 1027 ACCATTCGGCGTGAAGACACTTTCGATGACATGGGATTTTTCATCATCACTAGTTGAC 1086  
Qy 362 SerGlnAlaMetGlyAArgAlaGlyGlnValIleProArgThrTrpGlnThrAlaAapLyS 381  
Db 1087 TCTCAAGCTATAGGTCGTGTGGTGAAGTATCAACAGAACTTGGCAACAGCTGCAAA 1146  
Qy 382 AanaLySlyGlnPheGlyLySLeuProGluAAspGlyLyAAspAanAspAanPheAArgIle 401  
Db 1147 AACCAAAAAGATTTGGCGCTTGAAGAAGAAAGGCGATTAACCAAACTTCAGAAATC 1206  
Qy 402 LysArgTyIleSerLySlyThrIleAsnProAlaLeuThrHisGlyValaSerGlyTy 421  
Db 1207 AAACGCTACTGTCTAAATACACCATTAACCCAGCATCGCTCATGAGATTAGCGAGTAT 1266  
Qy 422 IleGlySerValGlnGlnGlyLySleAlaAspLeuValValTrpAsnProAlaPhePhe 441  
Db 1267 GTAGGTTCTGTAGAAATGGCAAAAGTGTGATTTGTGAGTGTGAGTCAAGATTCCTT 1326  
Qy 442 GlyValLyProLySleValIleLySlyGlyMetValValPheSerGlnMetGlyAap 461  
Db 1327 GCGGTGAACCCCAACATGATCATCAAGGTGGTTATTCATGATGATCAATGGCCGAT 1386  
Qy 462 SerAanaIleSerValProThrProGlnProValTyTyArgGlnMetPheGlyHisHIS 481  
Db 1387 GGGAAGCTTCTATCCCTACCCCAACAAGTTTATTAAGAGAAATGTTCCGCTCATCAT 1446  
Qy 482 GlyLyAlaAlaPheAspThrSerIleThrPheValSerLySlyValaLyTyTyGlnGly 501  
Db 1447 GGTAAGCCAAATATCATGCAACATCACTTTTGTCTTAAAGCGGCTTATGCAAAAGGC 1506  
Qy 502 ValLyGlnUtrLeuGlyLeuGlnAArgGlnValaIleProValLyAanCysAArgAanIle 521  
Db 1507 ATTAAGAAAGATTAAGGCTTGAAGACAAGTGTTCGCGTAAAAAATTCAGAAACATC 1566  
Qy 522 ThrLySlyAAspPheLySleAanAapLySThrAlaLySleThrValaAapProLySthr 541

Db 1567 ACTAAAAAGACATCAATTCAACGACACTACCGCTCATGTAAGTCAATCTGAAACT 1626  
Qy 542 PheGlnValPheValAAspGlyLySLeuCySThrSerLySProThrSerGlnAlaProLeu 561  
Db 1627 TACCATGTGTCGTGTGAATGGCAAAAGTAAGTACTTCTAAACCACTAAAGTGAAGCTTG 1686  
Qy 562 AlaGlnArgTyThrPhePhe 568  
Db 1687 GCGCAACTCTTATGACATTTTC 1707  
RESULT 13  
HECUREASE 2664 bp DNA linear BCT 03-MAY-1995  
LOCUS Helicobacter heilmannii urease, complete cds.  
DEFINITION L25079  
ACCESSION L25079  
VERSION L25079.1 GI:793908  
KEYWORDS urease.  
SOURCE Helicobacter heilmannii  
ORGANISM Helicobacter heilmannii  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Helicobacteraceae; Helicobacter.  
REFERENCE 1 (bases 1 to 2664)  
AUTHORS Solnick,J.V., O'Rourke,J., Lee,A. and Tompkins,L.S.  
TITLE Molecular analysis of urease genes from a newly identified  
uncultured species of Helicobacter  
JOURNAL Infect. Immun. 62 (5), 1631-1638 (1994)  
MEDLINE 9422523  
PUBMED 8168924  
COMMENT On May 3, 1995 this sequence version replaced gi:529422.  
Original source text: Helicobacter heilmannii (individual isolate  
2) DNA.  
FEATURES  
source location/Qualifiers  
1..2664  
/organism="Helicobacter heilmannii"  
/mol\_type="genomic DNA"  
/isolate="2"  
/db\_xref="taxon:35817"  
/lab\_host="BALB/c"  
200..204  
211..915  
/codon\_start=1  
/transl\_table=11  
/product="urease"  
/protein\_id="AA065722.1"  
/db\_xref="GI:529423"  
/translation="MKLRPEKLDKMLHYAGELAKORAKGKIKNYTEAVNLISAHW  
EARKGKSVADLNOEGRTLLKADVMFGVAHMHIEVGEIAGFPDGTGLVTHPTVEA  
GSDKLAPGEVILKNEIDITLNGKHAVOLKVKKRGDPVOVGSHPFPEVNKLADPRE  
KAYGRLLDIASTAVRFPGRBEKTVELIDIGNKRIVGFNMLVDRQADHDGKQALKR  
AKEKHFGTINCQDNK"  
919..923  
930..2636  
/codon\_start=1  
/transl\_table=11  
/product="urease"  
/protein\_id="AA065723.1"  
/db\_xref="GI:529424"  
/translation="MKKISRKEYVMGPTGDKVRLGDTDLLEVEHDCITTYGEIK  
FEGGKTRIDMGQINSBSHEDLVITNALVDVYGIYKADIGIKNGKIHGKAGNK  
DLQDVNCRLLCVGPATEMLAABGLVTAAGIDTHLHFSPOQIPFAFSGLITTMIGG  
TPADGTNAATLTTPGRNMLKEMLRASEBEYANLGLTGKGVNSFEPALIDOLEAGAGF  
KIHEDMGSTPEALINLALNIADKYDVQVALIHDTLNAECVDELTLEALIGRTIHTPTE  
GAGGHAARDVIMAGEFNILPASTNPTIPTKNTAEHMDLMCHLIDNKIKEDVEFA  
DSRIKPTIADKIDKIDMGISITSSDSQAMGRVGEVITRTMOTADKKKKEGRLPEE  
KQDNDFRIKIKITITNPALITGISTVSGSEVGYKADLVLMSPAFGKIKPMNLIK  
GGFIALSQMGANASIPTPOPYRREHGHGKAFDNIITFVSQAAENGKHEHGL  
QRVVLPRVKNCRNITKDKLKFNDVTAHIEVNPETYKVKVDGNEVTSAAADKLSLAOLYN  
LP"

ORIGIN  
Alignment Scores: 1.29e-143 Length: 2664  
Pred. No.: 2298.50 Matches: 418  
Score:



```

/transl table=11
/product="61 kDa protein"
/protein_id="CAA00663.1"
/db_xref="GI:41237"
/db_xref="GOA:P14917"
/db_xref="UniProt/Swiss-Prot:P14917"
/translation="MKKSRKEVSMWGPTGDDKRVLAGPTDLIAEVEHDYTIYGEELK
FGGGTIRREGMSQNNPSKEELDITLNLIVDTGYKADIGIDGKTAGIKGKGNK
DMQDVKNNLSVGRPALEALAGSLVTPAGITPHIHFIISPOQIPAPASGYTMIGGG
TGPADGTNATTTTPGRNLKMLRAEYSISNMGFLKGNASNDLSADQIEAGALGF
KIHEDWGTTPSAIINHALDVADKYDQVALIHTDLNEAGCVEDTMAIARQIMTTPHE
GAGGGHAPDIIRVAGEHNILPASTNPETPTVATEHMDMLVCHILDKSIEKVOF
ASRIRPOTIAEDLHDWGIETSTSSDQAMGRVEGVITRTQADKKKKEPKYKE
EKGDNDFRIKRYLSKTYTNPAAIHGISISYGVSEVGYKADLVMSAPFVGMPII
KCGTIALSOMGDANASITPTCPVYRHEMRAHKGAKYDANITPVQSAADKIKKEIG
LERQVLPVKNCRNITKDMQFNDITTAHIEVNPBTYHVPDGEVTSKPNKVSIAQLF
SIF"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 9 48e-144 Length: 1710
Score: 2297.00 Matches: 416
Percent Similarity: 86.77% Conservative: 76
Best Local Similarity: 73.37% Mismatches: 75
Query Match: 76.59% Indels: 0
DB: 6 Gaps: 0

```

US-09-904-994B-3 (1-568) x A07398 (1-1710)

```

QY 2 LysMetLysLysGlnGlnLysThrValAsnThrTyrgLysProThrLysGlyAspLysValArg 21
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 7 AAGATTAGCAAAAGAAATATGTTTCTATGATGCTTACTACTACAGCGCATTAAGTGAAGA 66
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 67 TTGGGGGATACAGACTGATCGCTGAAGTAGAACATGACTACCCATTATGCGCGAAGAG 126
QY 42 LeuLysPheGlyAlaGlyLysThrIleAspGlnGlyMetGlyGlnSerAsnSerProAsp 61
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 127 CTTAAATTTGGTGGCGGCTTAAACCTTAAGAGAGGAGATGAGCCAACTTAAACACCTTAC 186
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIleTyr 81
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 187 AAAGAAAGAGTGGATTATTAATCACTTAACGCTTATATCGTGATTAACCGGTATTATT 246
QY 82 LysAlaAspTlleGlyLysAsnGlyLysIleHisGlyTlleGlyLysAlaGlyAsnLys 101
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 247 AAAGCGGATATTGTTTAAAGATGCAAAATCGCTGCAATTGTTAAAGCGGTAAACAA 306
QY 102 AspMetGlnAspGlyValSerProHisMetValIleGlyValGlyThrGlnAlaLeuAla 121
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 307 GACATCCAAAGATGCGCTTAAACAAATCTTAGCGTAGGTCCTCTACTGAAGCCTTAGCC 366
QY 122 GlyGlnGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 367 GGGAAAGTTTGAATCGTAACGGCTGGTGGATTGACACACATCCACTTCAATTCACCC 426
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 427 CAACAAATCCCTTACACCTTTTCAGACGGGTGTAACAACATGATTTGGTGGTAACCGGT 486
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 487 CCTGCTGATGGCACTAATGCGACTACTTCACTCCGCGCAAGAAATTTTAAATGGAATG 546
QY 182 LeuArgAlaAlaGlnGlnLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 547 CTCAGAGCGCGCTGAACAATATCTATGAATTAAGTTTCTTGGCTAAAGTAAACGTTCT 606
QY 202 SerLysLysGlnLeuValGlnGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGly 221
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 607 AAGATGCGAGCTTACCGATCAAAATTGAACCGCGGTGCAATTGGCTTTAAATTCACGAA 666

```

```

QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyLysArg 241
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 667 GACTGGGCGACACACTCTTCTTGCATATCAATCATGGCTATAGATGTTCCGCGCAAAATGCA 726
QY 242 ValGlnValCysIleHisThrAspThrValAsnGlnAlaGlyTyrValAspAspThrLeu 261
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 727 GTCGAAGTCGCTATCCACACAGACACTTGAATGAAGCCGGTGTGTGTAACACATATG 786
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGlnGlyAlaGlyGlyHis 281
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 787 GCTGTATATTCGACGCACTATGACACTTTCACACTGAAGGCGCTGGCGCGGACAC 846
QY 282 SerProAspValIleThrMetAlaGlyGlnLeuAsnIleLeuProSerSerThrThrPro 301
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 847 GCTCTGATATATTAATTAAGTACCGCGTGAACACAACTTCTCCGCTTCCACTTAACCC 906
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetCysHis 321
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 907 ACCATCCCTTTCACCGCTGAATACAGACGACAGACATGACATGCTTATGCTGTGCCAC 966
QY 322 HisLeuAspLysArgIleArgGlnAspLeuGlnPheSerGlnSerArgIleArgProGly 341
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 967 CACTTGATTAATAAGATTAAGATGCGGTGAACACAACTTCTCCGCTTCCACTTAACCC 1026
QY 342 SerIleAlaAlaGlnAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1027 ACCATTGCGCGTGAAGACACTTTCATATGACATGGGGATTTTCTCATACCAAGTTCTGAC 1086
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProAspGlnThrTrpGlnThrAlaAspLys 381
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1087 TCTCAAGCGATGGCGCGTGGGTGAAGATTATCACTAGAACCTTGCACAAACACTACAA 1146
QY 382 AsnLysLysGlnPheGlyLysLeuProGlnLysAspAsnAspAsnPheArgIle 401
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1147 AACCAAGAAAGATTGGCCGCTTGAAGAAGAAAAAGCGATTAACGACAACTTCAGATC 1206
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGlyLys 421
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1207 AAACCTACTTCTTAATTAACCATTAACCCAGGAGATCCGCTCAAGGGATTAGCGAGTAT 1266
QY 422 IleGlySerValGlnGlnGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1267 GTAGGTTCAGTGAAGATGGGCAAGGTGCTGACTTGTGATGTGAGATCCACACTTCTT 1326
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValIlePheSerGlnMetGlyAsp 461
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1327 GCGTGAAGCCCAACATGATCAAAAGCGGATTCATTCGCTTAAGCCAAATGGCGCAT 1386
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGlnMetPheGlyHisHis 481
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1387 GCGAAGCTTCTATCTCCTACCCACACACCGGTTTATTAAGAGAAATGTTGCTCATCAT 1446
QY 482 GlyValAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGlnAsnGly 501
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1447 GGTAAAGCTAAATAGACATGCAACATCACTTGTGTCTCAAGCCGCTTATGACAAAGGC 1506
QY 502 ValLysGlnLysLeuGlnValLeuGlnArgGlnValLeuProValLysAsnCysArgAsnIle 521
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1507 ATTAAGAAAGATTAGACTTAAAGACAAGTGTGGCGGTAAATAATTCGAGAAATATC 1566
QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1567 ACTAATAAAGACATGCAATTCACACACTACTGCTCATCTGAAGTCAATCTGAACCT 1626
QY 542 PheGlnValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1627 TACCATGTGTTGTGATGCGCAAGAAAGTAACTTAAACCGACCAATAAAGTGAAGCTTG 1686
QY 562 AlaGlnArgTyrThrPhePhe 568
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 1687 GCGCACTCTTATGACATTTTC 1707

```

RESULT 15

A08818  
LOCUS A08818 1710 bp DNA linear PAT 09-AUG-1993  
DEFINITION C.jejuni DNA for 61 kDa protein.  
ACCESSION A08818  
VERSION A08818.1 GI:412247  
KEYWORDS  
SOURCE Campylobacter jejuni  
ORGANISM Campylobacter jejuni  
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;  
Campylobacteraceae; Campylobacter.  
1 (bases 1 to 1710)  
REFERENCE  
AUTHORS  
TITLE  
SEQUENCES OF NUCLEOTIDES CODING FOR A PROTEIN HAVING AN UREASIC  
ACTIVITY  
JOURNAL  
FEATURES  
source  
Patent: WO 9004030-A 8 19-APR-1990;  
Location/Qualifiers  
1. 1710  
/organism="Campylobacter jejuni"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:197"  
1. 1710  
/codon\_start=1  
/transl\_table=1  
/product="61 kDa protein"  
/protein\_id="CAA0881.1"  
/db\_xref="GI:412248"  
/db\_xref="GOA:P14917"  
/db\_xref="UniProt/Swiss-Prot:P14917"  
/translation="MKKISREYVSMYPTGDCVRLGDDLLAENEHDYTIYGEELK  
FGGKTLREGSGSNPSKEELDLITNALIVDTYGLKDIGKDKKIGGKAK  
DMQDGVKNNTSVGPAETALAGEGLIYTAGGIDPHIHFIHQOIPPTAASGVTTMGGG  
TGPAIDGNATTTTPGRNLKMLRAEYSNLCFLKGNASNDASLDQIEAATGP  
KIHEDWCTTSALNHALDVADKYDQVAIHDTLNEAGCVDVETMAALAGRTMHTFHE  
GAGGNAFDIIVKAGEHNIIPASTNPITPTVNTAEHMDLWCHLDKSIKEDVDF  
ADBRIRQPTIAEDTLDMGIFSTISDSQMGVGEVITRTWTQAKNKKEFGLKE  
EKGDNDNFRISKYLYKYTINPATAHIGSEVGVKADLVIMSPADPKGVKRMII  
KGGFIALSOKMGDANASTPTPOPVYRPMFPAHAKAKADANITTPSQAYDKGIKEELG  
LEKQVLPVKRCNRTKMDQFNDTTHIEVPEYTHVFDGKEVTSKPAKNSLALQLF  
SIF"

ORIGIN  
Alignment Scores:  
Pred. No.: 9.48e-144 Length: 1710  
Score: 2297.00 Matches: 416  
Percent Similarity: 86.77% Conservative: 76  
Best Local Similarity: 73.37% Mismatches: 75  
Query Match: 76.59% Indels: 0  
DB: Gaps: 0

US-09-904-994B-3 (1-568) x A08818 (1-1710)

QY 2 LysMetLysLeuGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21  
Db 7 AAGATTGACGAAAAAGAAATATGTTCTATGATGTCTCTACTACAGCGAATAAGTGAGA 66  
QY 22 LeuGlyAspThrAspLeuTTPAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41  
Db 67 TTGGGCGATTCAGACTTCGATCGCTGAAGTAGACATACATACCATTTATGGCGAAGAG 126  
QY 42 LeuLysPheGlyLysGlyLysThrIleArgGlnGluLysMetGlyGlnSerAspSerProAsp 61  
Db 127 CTTAATAATTCGGTGGCGGTAAACCCCTAAGAGAGAGCAATGAGCCAACTTAACAAACCCCTAGC 186  
QY 62 GluAsnThrLeuAspLeuValIleThrAspAlaMetIleIleAspTyrThrGlyIleTyr 81  
Db 187 AAGAAGAGAGTTGATTAATTAATACCTAACCTTTATTCGCGATTAACACCGGATTTAT 246  
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101  
Db 247 AAGCGGATTTGGTATTAAAGATGGCAAAATCCGTCGATTTGTTAAAGCGGCTTAACAAA 306  
QY 102 AspMetGlnAspGlyValIserProHisMetValGlyValGlyTyrThrGluAlaLeuAla 121

Db 307 GACATGCAAGATGCGCTTAAATAACATCTTAGCTGCTGCTACTGAAGCTTACGCC 366  
QY 122 GlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSerPro 141  
Db 367 GGTAAGATTGATTCGTAAAGCGGCTGGTGTATGACACACATCCATTCATTTACACCC 426  
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrHisPheGlyGlyIleThrGly 161  
Db 427 CAACAAATCCCTACAGCTTTTGGACGGGTGAACCAACATGATTTGGTGGTGAACCGGT 486  
QY 162 ProValAspGlyThrAsnAlaIleThrIleThrProGlyLysTyrLeuAsnLysValGlyMet 181  
Db 487 CTCTGATGCGACTAATGCGACTATGACTATCCACGAGAGAAATTTAAATGATGAT 546  
QY 182 LeuArgAlaAlaGluGluTyrSerMetAspValGlyPheLeuGlyLysGlyAsnSerSer 201  
Db 547 CTCAGACGGCTGAAGAAATTCATGAAATTTAGTTTCCTGGCTAAAGCTTACGCTTCT 606  
QY 202 SerTyrLysGlnLeuValGlnGlnValGluAlaGlyValIleGlyPheLysLeuHisGly 221  
Db 607 AACGATGCGAGCTTAGCGCGATCAAAATGAAGCCGGTGCATTCGCTTAAATTCACGAA 666  
QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTyrAsp 241  
Db 667 GACTGGGGCACCACTCTCTTCGATATCAATCAATCGCTTAATGTTGGCGACAAATACGAT 726  
QY 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261  
Db 727 GTGCAATCGCTATCCACACAGACACTTGAATGAAGCCGGTGTGTAAGAACACATATG 786  
QY 262 AsnAlaMetAsnGlyValArgAlaIleHisAlaTyrHisIleLeuGluValaGlyGlyHis 281  
Db 787 GCTGTAATTCGTGACCGCACTATGACCACTTTCCACACTAAGCGGTGCGGGACAC 846  
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301  
Db 847 GCTCTATATTTATTAAGTAGCGCGGGAACACAACTTTCTCCGCTTCCACTAACCCC 906  
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspPheLysMetThrCysHis 321  
Db 907 ACATCCCTTCCACCGGATACAGAAAGACAGACATGACATGCTTATGCTGTGCAC 966  
QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
Db 967 CACTTGGATTAAGCATTTAAAGAAAGATGTCAGTTCCCTATTCCAAAGATCCCGCTCAA 1026  
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361  
Db 1027 ACCATTCGGCTGAAGACACTTTCATGATGATGGGATTTTTCATACACAGTTCTGAC 1086  
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys 381  
Db 1087 TTCCAAGCATGAGCGCGGTGGGGAAGTATACATCACTGAACCTTGGCAACAGCTGACAA 1146  
QY 382 AsnLysLysGluLysPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401  
Db 1147 AACAGAAAGAAATTTGGCCCGCTTGAAGAAAGGCGCATGAACGCAACTTCAGGATC 1206  
QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValIleSerGluTyr 421  
Db 1207 AAAGCTACTTGTCTAATATCACCATTTAACCCGCGATCCCTCATGAGGATTAAGCAATAT 1266  
QY 422 IleGlySerValGlnGluGlyLysIleAlaAspLeuValIleTyrAsnProAlaPhePhe 441  
Db 1267 GTAGGTTCACTAAGAGAGGCAAGTGGCGTGAATGATGTGAGATCCAGACTTTT 1326  
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGlnMetGlyAsp 461  
Db 1327 GGGGTGAACCAACATGATCAATCAAGCGGATTCATTCGTTAAGCAAAATGGGGAT 1386  
QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGlnMetPheGlyHisHis 481  
Db 1387 GCGAAGCTTCTATCCCTACCCCAACCGGTTTATTAACAGAAATTTCCGCTCATCAT 1446

Qy 482 GlyValAlaIysPheAspThrSerIleThrPheValSerIysValAlaTyrGluAsnGly 501  
Db 1447 GGTAAAGCTAAATACATGCACAACATCATCTTTGTGTCTCAAGCGCTTATGCAAAAGGC 1506  
Qy 502 ValIysGluIysLeuGlyLeuGluArgGlnValLeuProValIysAsnCysArgAsnIle 521  
Db 1507 ATTAAGAAGAAATTAGCACTTGAAAGACAAGTGTGCCGTTAAATAATTGCAGAAATATC 1566  
Qy 522 ThrIysGluAspPheIysPheAsnAspIleThrAlaIysIleThrValAspProIysThr 541  
Db 1567 ACTAAAAAAGACATGCATTCAACGACACTACTGCTCACATTGAAGTCAATCCTGAACCT 1626  
Qy 542 PheGluValPheValAspGlyIysLeuCysThrSerIysProThrSerGlnValProIeu 561  
Db 1627 TACCATGTGTGTGATGCGAAAGAAAGTAACCTTAAACCAAGCCAAATAAGTGAGCTTG 1686  
Qy 562 AlaGlnArgTyrThrPhePhe 568  
Db 1687 GCGCAACTCTTTAGCATTTTC 1707

Search completed: February 15, 2005, 23:08:46  
Job time : 5478.47 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 15, 2005, 20:44:36 ; Search time 220.332 Seconds  
(without alignments)  
4218.194 Million cell updates/sec

Title: US-09-904-994B-3  
Perfect score: 2999  
Sequence: 1 MKMKKQEVYVVTGPTKGDV.....KLTCKPTSPQVPLAQRVTF 568

Scoring table:  
BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Xgapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 1202784 seqs, 818138359 residues  
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=x1h  
-Q=/cgn2\_1/USPTO.spool/US09904994/runat\_11022005\_145648\_25560/app\_query.fasta\_1.1102  
-DB=Issued\_Patents\_NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=p2n -NCM=ext HEADSIZE=500 -MITLEN=0 -MAXLEN=2000000000  
-USER=US09904994.@CGN\_1.1.77.@runat\_11022005\_145648\_25560 -NCPU=6 -ICPU=3  
-NO MMAP -LARGESUBSTRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2297	76.6	2735	2	US-08-920-095-1
2	2297	76.6	2735	2	PCT-US96-05800-1
3	2281	76.1	4824	4	US-09-431-705-1
4	2281	76.1	4824	4	US-09-431-705-1
5	2244	74.8	2619	2	US-08-467-822-19
6	2244	74.8	2619	2	US-08-467-822-19
7	2244	74.8	2619	3	US-08-466-248-19
8	2038.5	67.6	6131	1	US-07-732-242C-8
9	1828	61.0	1701	4	US-09-252-991A-10316
10	1826	60.9	1716	4	US-08-487-429A-3
11	1826	60.9	1716	5	PCT-US96-05320A-541
12	1826	60.9	1830121	4	US-09-557-884-1

13	1826	60.9	1830121	4	US-09-643-990A-1	Sequence 1, Appl
14	1821	60.7	1878	4	US-09-489-039A-2045	Sequence 2045, Ap
15	1818.5	60.6	1710	4	US-09-543-661A-1657	Sequence 1657, Ap
16	1818	60.6	1704	4	US-09-328-352-1786	Sequence 1786, Ap
17	1817	60.6	2400	1	US-08-967-513-1	Sequence 1, Appl
18	1817	60.6	2400	2	US-08-687-645B-1	Sequence 1, Appl
19	1809	60.3	87563	3	US-09-453-702B-57	Sequence 57, Appl
20	1808	60.3	8729	3	US-09-453-702B-258	Sequence 258, App
21	1789.5	59.7	1716	4	US-09-710-279-1828	Sequence 117, App
22	1789.5	59.7	1716	4	US-09-710-279-3828	Sequence 3828, Ap
23	1789.5	59.7	3234	4	US-09-710-279-3853	Sequence 3853, Ap
24	1788.5	59.6	1722	3	US-09-134-001C-2189	Sequence 2189, Ap
25	1784.5	59.5	5966	4	US-08-956-111E-22	Sequence 22, Appl
26	1784.5	59.5	5966	4	US-08-781-986A-122	Sequence 122, Appl
27	1761	58.7	1653	4	US-09-252-991A-10546	Sequence 10546, A
28	1708	57.0	1833	4	US-09-602-777A-13	Sequence 13, Appl
29	1626.5	54.2	1797	4	US-09-601-198-89	Sequence 89, Appl
30	1619	54.0	1625	4	US-09-602-777A-15	Sequence 15, Appl
31	1610	53.7	4403765	3	US-09-103-840A-2	Sequence 2, Appl
32	1610	53.7	4411529	3	US-09-103-840A-1	Sequence 1, Appl
33	1115	37.2	951	4	US-09-252-991A-10513	Sequence 10513, A
34	945.5	31.5	3475	4	US-09-710-279-4032	Sequence 4032, Ap
35	701	23.4	636	4	US-09-252-991A-10653	Sequence 10653, A
36	317	10.6	309	4	US-09-710-279-121	Sequence 121, App
37	315.5	10.5	292	4	US-09-313-294A-7157	Sequence 7157, Ap
38	189.5	6.3	1374	4	US-08-415-658-70	Sequence 20, Appl
39	175.5	5.9	2841	4	US-09-949-016-4735	Sequence 4735, Ap
40	175.5	5.9	2842	4	US-09-702-705-1800	Sequence 1800, Ap
41	175.5	5.9	2842	4	US-09-736-457-1800	Sequence 1800, Ap
42	175.5	5.8	1719	4	US-09-671-325-1800	Sequence 1800, Ap
43	173.5	5.8	1719	4	US-09-949-016-199	Sequence 199, App
44	173.5	5.8	4403	4	US-09-949-016-5632	Sequence 5632, Ap
45	173.5	5.8	5421	4	US-09-814-915A-66	Sequence 66, Appl

#### ALIGNMENTS

RESULT 1  
US-08-920-095-1  
; Sequence 1, Application US/08920095  
; Patent No. 5837240  
; GENERAL INFORMATION:  
; APPLICANT: Cynethia K. Lee et al.  
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,095  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,041  
; FILING DATE: 28-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 06132/013001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:



Alignment Scores:	
Pred. No.:	6,26e-247
Score:	2297.00
Percent Similarity:	86.77%
Best Local Similarity:	72.33%
Query Match:	76.59%
DB:	5
Length:	2735
Matches:	416
Conservative:	76
Mismatches:	75
Indels:	0
Gaps:	0

Dd	1248	CAACAATACCTTACAGCTTTTTCGAAAGCGGTGTAACACCATATGTTGGTGGAAACCGGT	130
Qy	162	ProValaAerGIYThrAsnAlaThrThrIleThrProGlyIleTyrPAsnLeuH1sArgMet	181
Dd	1308	CTGTGTGATGGCAGTAATATGCGCATCTACTATCACTCCAGGCGAAGAAATTTTAAATGGATG	136
Qy	182	LeuArgAlaAlaGluGluTyrSerMetCAsnValGlyPheLeuGlyLysGlyAsnSerSer	201
Dd	1368	CTCAAGCGCGGTGAAGAAATATCTTATGTAATTAAGTTCTTGGCTTAAAGTAAACGCTTCT	142
Qy	202	SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaLeuGlyPheLysLeuH1sGlu	221
Dd	1428	AACGATGCGAGCTTACGCCGATTAATGTAAGCGGCGCATTTGGCTTTAAATTCACGAA	148
Qy	222	AspTrpGIYThrThrProSerAlaIleAspH1sCysLysSerValAlaAspGluTyrAsp	241
Dd	1488	GACTGGGGCACAACCTCTTCTTGCATCAATCATGCGTTAGATGTTGGCGGACAAATTCGAT	154
Qy	242	ValGlnValCysIleH1sThrAspThrValaAsnGluAlaGlyTyrValaAspAspThrLeu	261
Dd	1548	GTGCAGATGCGGTATGCACACAGACACTTTTGAATGAAGCGCGTGTGTGAAGACATATAG	160
Qy	262	AsnAlaMetCAsnGlyArgAlaIleH1sAlaTyrH1sIleGluGlyAlaGlyLysGlyH1s	281
Dd	1608	GCTGTATATGCTGACGACGACTATGACACTTTCCACTGAAGGCGCTGGCGCGGACAC	166
Qy	282	SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrTyrPro	301
Dd	1688	GCTCTGATATTTATTAAGATAGCGCGGTGACACAACTTTCTCCGCTTCCACTAACCC	172
Qy	302	ThrIleProTyrThrIleAsnThrValaAlaGluH1sLeuAspMetLeuMetThrCysH1s	321
Dd	1728	ACCATCCCTTCAACCGTGAAATACAGAACACAGACCATGGAATCATGCTTATGGTGTGCCAC	178
Qy	322	H1sLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly	341
Dd	1788	CACTTGGATTAAGCATTAAGAAAGAAATGTTCACTTGCCTGATTCAGAGATCCGCGCTCAA	184
Qy	342	SerIleAlaAlaGluAspValLeuH1sAspMetGlyValIleAlaMetThrSerSerAsp	361
Dd	1848	ACCATGGCGGTGAAGACACTTTGCATGACATGGGATTTTCTCAATCACAGTTCTGAC	190
Qy	362	SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrGlnThrAlaAspLys	381
Dd	1908	TCTCAAGCGATGGCGCGTGTGGGTGAATGTTATCACTAAGACTTGGCAACACGTGCATA	196
Qy	382	AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle	401
Dd	1968	AACAGAAAGAAATTTGGCGCGCTTGAAAGAAAGAAAGCGATMAAGACAACTTCAGGATC	202
Qy	402	LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrH1sGlyValSerGluTyr	421
Dd	2028	AAAGCGTACTGTCTTAATATACACACTTAACCCAGGATCGCTCAATGGGATTTACGAGAT	208
Qy	422	IleGlySerValGluGluGlyLysIleAlaAspLeuValValTyrAsnProAlaPhePhe	441
Dd	2088	GTAGGTTCAATGAAGGTGGGCAAGGTGCTGACTTGATATGTGAGATGCACACTTCTTT	214
Qy	442	GlyValLysProLysIleValIleLysGlyGlyMetValLysSerGluMetGlyAsp	461
Dd	2148	GGCGGAAACCAACATCATCATCAACAAAGCGGATTCATTCGCTTAAAGCAAAATGGCGCAT	220
Qy	462	SerAsnAlaSerAlaProThrProGlnProValTyrTyrArgGluMetPheGlyH1sH1s	481
Dd	2208	GCGAAGCTTCTATATCCCTTACCCACAAACCGGTTTATTCACAGAAATGTTCCCTCATAT	226
Qy	482	GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly	501
Dd	2268	GTAAAGCTTAATATGATGCAAAACATCACTTTTGTGTCTCAAGCGGCTTATGACAAAGCG	232
Qy	502	ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnIle	521

Db 2338 ATTAAGAGAAATTAAGACTTGAAAGACAAGTGTCCGCTAAATAATGTCAGAAATATC 2387  
QY 522 ThrLVsAspPheLVsPheAsnAspLVsThrAlaLVsIleThrValAspProLVsThr 541  
Db 2388 ACTTAATAAAAGACATGCAATTCACGACACTGCTCACTGCAATTCGAACTGAACT 2447  
QY 542 PheGluValPheValAspGlyLVsLeuCySerSerLVsProThrSerGluValProLeu 561  
Db 2448 TACCATGTGTTGCTGGATGGCAAGAAAGTAACCTTAAACCAAGCCAAATAAGTAGACTTG 2507  
QY 562 AlaGluArgTyrThrPhePhe 568  
Db 2508 GCGCAACTCTTATGCACTTTC 2528

## RESULT 3

US-09-431-705-1  
Sequence 1, Application US/09431705  
Patent No. 6585975  
GENERAL INFORMATION:  
APPLICANT: Kleantinous, Harold  
APPLICANT: Londono-Arcila, Patricia  
TITLE OF INVENTION: Use of salmonella vectors for  
FILE REFERENCE: 06132/060001  
CURRENT APPLICATION NUMBER: US/09/431,705  
NUMBER OF SEQ ID NOS: 52  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 4824  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: includes sequences from Helicobacter pylori,  
OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli  
NAME/KEY: CDS  
LOCATION: (2) ... (31)  
NAME/KEY: CDS  
LOCATION: (41) ... (61)  
NAME/KEY: CDS  
LOCATION: (65) ... (799)  
NAME/KEY: CDS  
LOCATION: (803) ... (2512)  
NAME/KEY: CDS  
LOCATION: (2516) ... (2692)  
NAME/KEY: CDS  
LOCATION: (2696) ... (2896)  
NAME/KEY: CDS  
LOCATION: (2900) ... (3322)  
NAME/KEY: CDS  
LOCATION: (3326) ... (3385)  
NAME/KEY: CDS  
LOCATION: (3389) ... (3406)  
NAME/KEY: CDS  
LOCATION: (3410) ... (3466)  
NAME/KEY: CDS  
LOCATION: (3470) ... (3598)  
NAME/KEY: CDS  
LOCATION: (3602) ... (3661)  
NAME/KEY: CDS  
LOCATION: (3665) ... (3697)  
NAME/KEY: CDS  
LOCATION: (3701) ... (3769)  
NAME/KEY: CDS  
LOCATION: (3773) ... (3817)  
NAME/KEY: CDS  
LOCATION: (3821) ... (3844)  
NAME/KEY: CDS  
LOCATION: (3848) ... (3889)  
US-09-431-705-1

Alignment Scores:

Pred. No.: 9,78e-245 Length: 4824  
Score: 2281.00 Matches: 414  
Percent Similarity: 86.42% Conservative: 76  
Best Local Similarity: 73.02% Mismatches: 77  
Query Match: 76.06% Indels: 0  
Gaps: 0  
DB: 4

US-09-904-994B-3 (1-568) x US-09-431-705-1 (1-4824)

QY 2 LysMetLVsLVsGluGluThrValAsnThrTyrGlyProThrLVsGlyValAspValArg 21  
Db 812 AAGATTAGCAGAAAGAAATGTTCTATGATGGTCTTACTACAGCGCATTAAGTGACA 871  
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTyrThrTyrGlyGluGlu 41  
Db 872 TTGGGCGATACAGACTTGATCGCTGAAGTAACAATGACTACACCATTTATGGCGAAGG 931  
QY 42 LeuLVsPheGlyAlaGlyLVsThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61  
Db 932 CTAAATTCGGTGGCGGTAAACCTTAAGAGAGAGCATGACCAATCTAACAACCTTACG 991  
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleAspTyrThrGlyIleTyr 81  
Db 992 AAGAGAGCTTGATTTAATTAATTAATCACTTAACGCTTAAATGCTGATTAACCGGATTTAT 1051  
QY 82 LysAlaAspIleGlyIleLVsAsnGlyLVsIleHisGlyIleGlyLVsAlaGlyAsnLVs 101  
Db 1052 AAGCGGATTTGGTATTTAAATGCAAAATCGCTGGCATTTGTTAAAGCGGTAAACAA 1111  
QY 102 AspMetGluAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeuAla 121  
Db 1112 GACATGCAAGATGGCGGTAAACCAATCTTACCGTCTTACGTAACCTTAAGCC 1171  
QY 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141  
Db 1172 GGTGAAGGTTTATGATGTAACGGCTGTGTGTATTAACAACAATCCACTTCACTTCAACC 1231  
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161  
Db 1232 CAACAAATCCCTACAGCTTTTCCAAAGCGGTGTAACAACATGATTTGGTGGAAACGGT 1291  
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLVsTrpAsnLeuHisArgMet 181  
Db 1292 CCTGCTGATGGCATAATGCACTACTACTCATCCAGGCAAGAAATTTAAATGGAAG 1351  
QY 182 LeuArgAlaAlaGluGluThrSerMetAsnValGlyPheLeuGlyLVsGlyAsnSerSer 201  
Db 1352 CTCAGAGCGGCTGAAGAAATATTCTATGAATTAAGTTCTTGGCTTAAGTAACGCTTC 1411  
QY 202 SerLVsLVsGluLeuValGluGluValGluAlaGlyAlaIleGlyPheLVsLeuHisGlu 221  
Db 1412 AACGATGCGAGCTTGAAGCGCATGAATTAAGCGGATTTGGCTTTCATTTACGAA 1471  
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241  
Db 1472 GACTGGGGCACCACTCTTCTCAATCAATCAATGCGTTGAGATGTTGGCGCAAAATACGAT 1531  
QY 242 ValGluValCysIleHisThrAspThrValAsnGluValGlyTyrValAspAspThrLeu 261  
Db 1532 GTGCAAGTCTATGCGCACACACACTTGAATGAAGACCGGTGTGTGAACAACACTATG 1591  
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGluGlyValGlyGlyGlyHis 281  
Db 1592 GCTGCTATTTGCGACGCACTATGACACTTTCACACATGAAGGCGCTGCGCGGACAC 1651  
QY 282 SerProAspValIleThrMetAlaGlyLeuAsnIleLeuProSerSerThrThrPro 301  
Db 1652 GCTCTGATATTAATTAAGTAGCCGATGAACAACAATCTTCCGCTTCCACTTAACCCC 1711  
QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321  
Db 1712 ACCATCCCTTTTCAACCGTGAATACAGAAAGACAGACATGACATGATGTTATGTTGTCAC 1771

```
Qy 322 HistLeuApLySArgIleArgIleuApLeuGlnPheSerGlnSerArgIleArgProGly 341
Db 1772 CACTTGGATTAAAGACATTAAAGAAAGATGTTCACTTCCTGCTTCAAGATTCGGCTTCAA 1831
Qy 342 SerIleAlaIaGluApValIleuHisApMetGlyValIleAlaMetThrSerSerAsp 361
Db 1832 ACCATTGGCGCTGAAGACATTTGATGATGATGGGATTTTCTCAATCACCAGTTCTGAC 1891
Qy 362 SerGlnAlaMetGlyValGAlaGlyValIleProAlaGlnThrTrpGlnThrAlaApLyS 381
Db 1892 TCTCAACGATGGGCGCGTGGTGGTGAAGTATCTAGAACCTTGGCAACGCTGACCAA 1951
Qy 382 AsnLySLeuGluPheGlyLysLeuProGluApGlyLysAspAsnAspPheArgIle 401
Db 1952 AACAGAAAGAAATTTGGCGCTTGAAGAGAGAAAAAGCGATTAACGACAACTTCAGATC 2011
Qy 402 LysArgTrpIleSerLysTrpThrIleAsnProAlaLeuThrHisGlyValSerGluTrp 421
Db 2012 AACCGTACTGTTGCTAAATACACCATTAACCAAGGATCGCTCATGGGATTAAGGAGTAT 2071
Qy 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValIleTPAsnProAlaPhePhe 441
Db 2072 GTAGGTTCAAGTAGAAGAGCGGCAAGTGGCTGATTTGTTGAGAGTCCAGCAATTCCTT 2131
Qy 442 GlyValLysProLysIleValIleLysGlyLysMetValIlePheSerGluMetGlyAsp 461
Db 2132 GGCGTGAACCCAAACATGATCATCAAGCGGATTCATTGCTTAAGCCAAATGGCGCAT 2191
Qy 462 SerAsnAlaSerValProThrProGlnProValTrpTrpArgGluMetPheGlyHisHis 481
Db 2192 GCGAAGCGTTTATCCCTACCCCAACACCGGTTTATTAACAGAAATTTGGCTCATCAT 2251
Qy 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTrpGluAsnGly 501
Db 2252 GGTAAGCTAAATACGATGCAAGCATCTTTGTGTCTCAAGCGGCTTAAGCAAAAGGC 2311
Qy 502 ValLysGluLysLeuGlyLysGluAlaGlnValIleuProValLysAsnCyArgAsnIle 521
Db 2312 ATTAAAGAGATTAAGACCTTGAAGACAGATGTTGGCGTAAATAATTTGCAAAATATC 2371
Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2372 ACTAAAGAAAGACATGCAATTCAACGACACTACCGTCCATTTGAAGTAACTCTGAAACT 2431
Qy 542 PheGluValPheValAspGlyLysLeuCyThrSerLysProThrSerGlnValProLeu 561
Db 2432 TACATATGTTCTGGATGGCAAGAGTAACTTCTAAACAGCAATTAAGTGAAGCTTG 2491
Qy 562 AlaGlnArgTrpThrPhePhe 568
Db 2492 GCGCAACTCTTAAACATTTTC 2512

RESULT 4
US-09-431-705-19
; Sequence 19. Application US/09431705
; Patent No. 6585875
; GENERAL INFORMATION:
; APPLICANT: Kleantchou, Harold
; APPLICANT: Londono-Arcila, Patricia
; APPLICANT: Freeman, Donna
; TITLE OF INVENTION: Use of salmonella vectors for
; TITLE OF INVENTION: vaccination against helicobacter infection
; FILE REFERENCE: 06132/060001
; CURRENT APPLICATION NUMBER: US/09/431.705
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 4824
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: includes sequences from Helicobacter pylori,
```

```
OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
; NAME/KEY: CDS
; LOCATION: (3893) ... (3934)
; NAME/KEY: CDS
; LOCATION: (3938) ... (4027)
; NAME/KEY: CDS
; LOCATION: (4031) ... (4285)
; NAME/KEY: CDS
; LOCATION: (4289) ... (4300)
; NAME/KEY: CDS
; LOCATION: (4304) ... (4408)
; NAME/KEY: CDS
; LOCATION: (4412) ... (4471)
; NAME/KEY: CDS
; LOCATION: (4475) ... (4588)
; NAME/KEY: CDS
; LOCATION: (4592) ... (4669)
; NAME/KEY: CDS
; LOCATION: (4673) ... (4711)
; NAME/KEY: CDS
; LOCATION: (4715) ... (4774)
; NAME/KEY: CDS
; LOCATION: (4784) ... (4824)

US-09-431-705-19

Alignment Scores:
Pred. No.: 9.78e-245 Length: 4824
Score: 2281.00 Matches: 414
Percent Similarity: 86.42% Conservative: 76
Best Local Similarity: 73.02% Mismatches: 77
Query Match: 76.06% Indels: 0
DB: Gaps: 0

US-09-904-994B-3 (1-568) x US-09-431-705-19 (1-4824)
Qy 2 LysMetLysLysGluTrpValAsnThrTrpGlyProThrLysGlyAspLysValArg 21
Db 812 AAGATTACGCAAAAGAAATATGTTCTATGTATGCTCTTACAGGCGATTAAGTGA 871
Qy 22 LeuGlyAspThrAspLeuTPAlaGluValGluHisAspTrpThrTrpGlyGluGlu 41
Db 872 TTGGCGCATACAGACTTGATGCTGAAGTAGAATGATGATCAATTTATGGGAAG 931
Qy 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61
Db 932 CTTAAATTCGGTGGCGGTAAACCTTAAGAGAGCATGAGCCATCTAACACACCTTAC 991
Qy 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTrpThrGlyIleTrp 81
Db 992 AAAGAGAGTGGATTTAATTATGATGATGCTTTAATCGGATTAACCGGATTTAT 1051
Qy 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
Db 1052 AAAGCGATATTTGTTATTAAGATGGCAAAATCGCTGGCATTTGTTAAAGCGGTAA 1111
Qy 102 AspMetGlnAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeuAla 121
Db 1112 GACATGCAAGTGGCGGTAAACCAATCTTAAGCTTACCTGCTAGTGAAGCCTTAC 1171
Qy 122 GlnGluGlyMetIleIleThrAlaGlyLysLeuSerHisThrHisPheLeuSerPro 141
Db 1172 GGTAAAGTGGTGAATTAAGTGAAGCTGGTGGTATTAACAACATCACTTCAATTCACC 1231
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThrGly 161
Db 1232 CAACAAATCCCTTAAGCTTTTGAAGCGGTGAACAACATGATGGTGGTGAACCGGT 1291
Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisAspMet 181
Db 1292 CTTGCTGATGGCACTAATGCGACATCACTTCACTCCAGAGAAATAATTTAAATGATG 1351
Qy 182 LeuArgAlaIaGluGluTrpSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
```



Percent Similarity: 85.01%  
 Best Local Similarity: 72.49%  
 Query Match: 74.82%  
 DB: 2  
 Gaps: 0

US-09-904-994B-3 (1-568) x US-08-467-822-19 (1-2619)

Conservative: 71  
 Mismatches: 85  
 Indels: 0

2 LysMetLysbSGInLgUlyrValAsnThrTyrGlyProThrLysGlyAspLysValArg 21  
 772 AAGATTTCACGAAAAAGATATGTTTCTATGTAATGCTCCACTACCGGGAGATGTTTGA 831  
 22 LeuGlyAspThrAspLeuTPDAlaGluValAspTyrThrTyrGlyGlu 41  
 832 CTCGGCGACATGATTTGATCTTGAAGTGGAGCATGATGGACACCTTAATGGTGAAG 891  
 42 LeuLysPheGlyAlaGlyLysThrLLeaRgGluGlyMetGlyGlnSerAsnSerProAsp 61  
 892 ATCAAAATTTGGGGCGGTAAACCTATCCGATGGATGGATCAAAACCAATAGCCCTAGC 951  
 62 GluAsnThrLeuAspLeuValLLeaThrAsnAlaMetLLeaAspTyrThrGlyLLeaTyr 81  
 952 TCTTATGAATTAATGATTTGCTGCTACTTAACCCCTCATTTGGAGCTATACGGCATTTAC 1011  
 82 LysAlaAspLLeaGlyLLeaLysAsnGlyLysLLeaGlyLLeaGlyLysAlaGlyAsnLys 101  
 1012 AAAGCCGACATTTGGGATTAAGACGCAAGATTTGAGCATTTGGCAAGGCGCAATAG 1071  
 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeuAla 121  
 1072 GACATGCAAGATGCGTATGATTAATATCTTGGCTGAGCTCTGCTACAGAGGCTTTGGCA 1131  
 122 GlyLysGlyMetLLeaThrAlaGlyLLeaAspSerHisThrHisPheLeuSerPro 141  
 1132 GCTAGAGGCTTGAATTTGATGACCGCTGGTGCATCGATACGATATTCATCTTATCTCTCC 1191  
 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyGlyThrGly 161  
 1192 CAACAAATCCCTACTGCTTTGCCAGCGGGTTACAAACATGATTTGGAGGAGGCAAGCA 1251  
 162 ProValAspGlyLysAsnAlaThrThrLLeaProGlyLysTPDAsnLeuHisArgMet 181  
 1252 CCTCGAGTGGCAGAAATGCGACCAACCATCTCCCGGACCGCTTAATCTTAAAAAGTGTG 1311  
 182 LeuArgAlaAlaGluGlyLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201  
 1312 TTGGGTGAGCGCAAGATACGCGCATGATCTTGGCTTTGGCTTAAGGGGAATGTGCT 1371  
 202 SerLysLysGlnLeuValGluGlnValAlaGlyAlaGlyLLeaGlyPheLysLeuHisGlu 221  
 1372 TAGCAACCCCTTTACCGGATCAGATTGAACAGCGGGCATTTGGTTTAAAAATCCAGCA 1431  
 222 AspTyrGlyLysThrProSerAlaLLeaAspHisCysLeuSerValAlaAspGlyLysAsp 241  
 1432 GACTGGGGGAAGCACACCTGACGATTTACCACTGCTCATATGTCGCGAGTAATACGAT 1491  
 242 ValGlnValCysLLeaThrAspThrValAsnGluAlaGlyLysValAspAspThrLeu 261  
 1492 GTGCAAGGCTTATCCACACCGATACCTTAAACAGAGGGGCTGTGTGAAGAACACCTTA 1551  
 262 AsnAlaMetAsnGlyLysAlaLLeaHisAlaLysLLeaGlyLysAlaGlyLysGlyHis 281  
 1552 GAGCGCATTTGGCGGGCGGACATCATCTTCCACTGAAGGGGCTGGGGGTGGACAC 1611  
 282 SerProAspValLLeaThrMetAlaGlyLLeuLeuAsnLLeuProSerSerThrPro 301  
 1612 GCTCCACATGTTATCAAAATGTGCGAGGAATTTAAACATTTACCGGCTCTTAAACCGG 1671  
 302 ThrLLeaProTyrThrLLeaThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321  
 1672 ACCATTCCTTTCACCAAAACACGAAACCGACGACATGACATGTAATGATGTGGCAC 1731  
 322 HisLeuAspLysArgLLeaGlyLysLeuGlnPheSerGlnSerArgLLeaArgProGly 341

1732 CACTTGGATAAAGATATCAAGAAAGATGTCAGTTTGGCCATTCGAGGATTCGCCCCAA 1791  
 342 SerLLeaAlaAlaGluAspValLeuHisAspMetGlyValLLeaHisMetThrSerSerAsp 361  
 1792 ACTATCGGGCTGAAGCAACCTCCATGACATGGAGATCTTTTCTATACCAAGCTCCGAC 1851  
 362 SerGlnAlaMetGlyLysAlaGlyLLeaValLLeaProArgThrTPDAlaAspLys 381  
 1852 TCTAGGCTATGGAACCGCTGAGGAGTGTATCACACGACCTTGGCAGACAGACAAA 1911  
 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgLLe 401  
 1912 AACAAAAAGATTTGGCGGCTTGAAGAGGAAAAAGCCGCTTAACGACAACTTCGCGATC 1971  
 402 LysArgTyrLLeaSerLysTyrThrLLeaAsnProAlaLeuThrHisGlyValSerGlyLys 421  
 1972 AAACGCTACATCTTAAATACACATACACCCCGGATCCGGCATGGGATTTGACATAT 2031  
 422 LLeaSerValGluGluGlyLysLLeaAspLeuValLLeaAspLeuValLLeaAsnProAlaPhePhe 441  
 2032 GTGGGCTCTGAGAAAGTGGCAATACGCGACCTCGTGTGGAGTCCGGCTTTCTTT 2091  
 442 GlyValLysProLysLLeaValLLeaLysGlyLysMetValAlaPheSerGluMetGlyAsp 461  
 2092 GGCAATTAAGCCCATATGATTAATTAAGGCGGATTTATGCGCTCTCAAAATGGCGCAT 2151  
 462 SerAsnAlaSerValProThrProGlnProValLysTyrArgGluMetPheGlyHis 481  
 2152 GCGAATGCTTATTTCCACCCCTCACCCGCTCTATTAACCGTAAAGTGTGGACACAT 2211  
 482 GlyLysAlaLysPheAspThrSerLLeaThrPheValSerLysValAlaTyrGluAsnGly 501  
 2212 GGGAAAAACAAATTCGACACCAATATCATCTTGTGTCGCAAGGGCTTTACAGCAGG 2271  
 2272 ATCAAGAAAGAACTAGAGGCTATGATCGCGCGCACCGCAAGTGAAGAACTGTGCAATATC 2331  
 502 ValLysGluLysLeuGlyLysGluArgGlnValLeuProValLysAsnCysArgAsnLLe 521  
 2332 ACTAAAAAGGCTCAAAATTCACAGATGTGACCCGACATATGATGTCACACCTGAAACC 2391  
 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysLLeaThrValAspProLysThr 541  
 2392 TATAGGTGAAGATGATGGCAAGAGATTAACCTTAAAGCAGCATGAATTAAGCTTA 2451  
 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561  
 2452 GCGCACTTATATATTTGTTT 2472

RESULT 6  
 US-08-432-697-19  
 Sequence 19, Application US/08432697  
 Patent No. 6248330  
 GENERAL INFORMATION:  
 APPLICANT: Labigne, Agnes  
 APPLICANT: Sauerbaum, Sebastian  
 APPLICANT: Ferrero, Richard L.  
 APPLICANT: Thibierge, Jean-Michel  
 TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
 TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
 TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
 NUMBER OF SEQUENCES: 44  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &  
 STREET: 1300 I Street, N.W.  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005-3315  
 COMPUTER READABLE FORM:

```

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0137-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 2619 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc_feature
LOCATION: 31..36
OTHER INFORMATION: /standard_name="Shine-Dalgarno"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 756..759
OTHER INFORMATION: /standard_name="Shine-Dalgarno"
OTHER INFORMATION: sequence."
US-08-432-697-19

```

```

Alignment Scores:
Pred. No.: 5.14e-241 Length: 2619
Score: 2244.00 Matches: 411
Percent Similarity: 85.01% Conservative: 71
Best Local Similarity: 72.49% Mismatches: 85
Query Match: 74.82% Indels: 0
DB: 3 Gaps: 0

```

US-09-904-994b-3 (1-568) x US-08-432-697-19 (1-2619)

```

QY 2 LysMetLysLysGlnGlyTrpValaenThrTrpGlyProthrLysGlyAspLysValArg 21
DB 772 AAGATTTCAGAAAGAAATAGTTTCATGTATGTGCCACTACCGGGATCGTTAG 831
QY 22 LeuGlyAspThrAspLeuTrpAlaGluValGluHisAspTrpThrTrpGlyGluGlu 41
DB 832 CTCGGGAGACCTGATTGATCTTAAGAGTGGACATGATTCGACCACTTATGCTGAAG 891
QY 42 LeuLysPheGlyAlaGlyLysThrIleArgGlnGlyMetGlyGlnSerAsnSerProAsp 61
DB 892 ATCAAAATTTGGGGCGGTAAACATATCCGTGATGGATGAGTCAAAACAATATACCTTAC 951
QY 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTrpThrGlyIleTrp 81
DB 952 TCTTATGAATTAATTTGGTCTCACTACGACCTCATGTTGAGCATTAACGGGCATTTAC 1011
QY 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
DB 1012 AAAGCGGACACTTGGATTAAAGACGGCAAGATTGCAAGCATTTGCAAGCGACATTAAG 1071
QY 102 AspMetGlnAspGlyValSerProHisMetValValGlyValGlyLysTrpGluAlaLeuAla 121
DB 1072 GACATCAAGATGGCTAGATTAATCTTTGGTGTGCTCTCTACAGAGGCTTTGGCA 1131
QY 122 GlyGlnGlyMetIleIleThrAlaGlyLysIleAspSerHisThrHisPheLeuSerPro 141
DB 1132 GCTGAGGCTTGTATTAACCGCTGGTGGCATGATTAATCACTTATCTCTCCC 1191
QY 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyLysThrGly 161

```

```

DB 1192 CAACAATCCCTTACTGCTTTTCCAGCGGGTTTACAAACATATTTGAGAGGACAGAA 1251
QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrpAsnLeuHisArgMet 181
DB 1252 CCGGGGATGGACAGAAATCGACACACATCACTCCCGGAGGGCTTATTTAAAGATATG 1311
QY 182 LeuArgAlaAlaGlyGluTrpSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201
DB 1312 TTGGGTGCGACCGGAAGAAATAGCCATGAATCTAAGCTTTTGGCTTAAGGGAATGTCT 1371
QY 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221
DB 1372 TAGCAACCTCTTTACGCATCAGATTGAAGACGGGGCATTTGGTTTAAATCCACGAA 1431
QY 222 AspTrpGlyThrThrProSerAlaIleAspHisCysLysSerValAlaAspGlyTrpAsp 241
DB 1432 GACTGGGGAAGACACCTGCACTATTCACCACTGCTCAATGTGCGCATGAATACGAT 1491
QY 242 ValGlnValCysIleHisThrAspThrValaenGluAlaGlyTrpValaAspAspThrLeu 261
DB 1492 GTGCAAGTGGCTATCCACACGATACCTTAACGAGGGGGCTGTGTAAAGACACCTTA 1551
QY 262 AsnAlaMetAsnGlyArgAlaIleHisAlaTrpHisIleGlnGlyAlaGlyLysGlyHis 281
DB 1552 GAGCGATTCGCGGGCGCACATCCATCTTCCACACTGAAGGGGCTTGGGGGTGGACAC 1611
QY 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrPro 301
DB 1612 GCTCCAGATGTATCAAAATGGACGGGGAATTTAATCAATCTACCGCTCTACTAACCG 1671
QY 302 ThrIleProTrpThrIleAsnThrValaGluHisLysAspMetLeuMetThrCysHis 321
DB 1672 ACCATTCCTTACCAAAAACACTGAAGCGGACATGAGCATGTATATGTGTGTGCCAC 1731
QY 322 HisLysAspLysArgIleArgGlnAspLeuGlnPheSerGlnSerArgIleArgProGly 341
DB 1732 CACTTGATTAAGATATCAAGAGATGTGCAATTGCGGATTCGAGATTCGAGTCCGCCAA 1791
QY 342 SerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAlaMetThrSerSerAsp 361
DB 1792 ACTATCGCGGTGAAGACCACTCACTGACATGGGATTTTCTTATCCACAGCTCCGAC 1851
QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTrpGlnThrAlaAspLys 381
DB 1852 TCTCAGCGCTATGGAGCGCTTACGCGATGATCACACGACCTTGGAGACACAGACAA 1911
QY 382 AsnLysLysGlnPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401
DB 1912 AACAAAAGAGATTTGGGCGCTTGAAGAGGAAAAAGCGGATTAACGAACTTCCGCATC 1971
QY 402 LysArgTrpIleSerLysThrIleAsnProAlaLeuThrHisGlyValaSerGlyTrp 421
DB 1972 AAACGCTATCATCTTAATAATACCAATCAACCCCGGAGTGGCGCATGTTCTCATAT 2031
QY 422 IleGlySerValGlnGlyLysIleAlaAspLeuValaValTrpAsnProAlaPhePhe 441
DB 2032 GTGGGCTCTGTGGAACTGGGCAAAATACGCGACCTCTGTCTTGGAGTCCGGCTTCTT 2091
QY 442 GlyValLysProLysIleValIleLysGlyGlyMetValaValPheSerGluMetGlyAsp 461
DB 2092 GGCATTAAGCCCAATATGATTAATTAAGGGCGGATTAATGACCTCTCTCAAAATGGCGCAT 2151
QY 462 SerAsnAlaSerValProThrProGlnProValIleTrpArgGluMetPheGlyHisHis 481
DB 2152 GCCAATGCTCTTATCCACCCCTCAGCCGCTATTAACGGGAATGTTTGGACCCAT 2211
QY 482 GlyValAlaLysPheAspThrSerIleThrPheValaSerLysValaAlaTrpGluAsnGly 501
DB 2212 GGGAAAAAATAATTTGACACCAATATCACTTCTGTCTCCCAAGCGGCTTAACAAGGAGG 2271
QY 502 ValLysGluLysLeuGlyLysGlnArgGlnValaLeuProValaLysAsnCysArgAsnIle 521

```



Db 2272 ATCAAGAGAACTAGGCTAGATCGCGCGACCGCAGTGAAGAACTGTGCAATATC 2331  
Qy 522 ThrlyslAspPhePheAsnAspLysThrAlaLysIleThrValAspProLysThr 541  
Db 2332 ACTTAAAGAGACCTCAATTAACGATGTGACCGACATATTGATGTAACCTGAAC 2391  
Qy 542 PheGluValPheValAspGlyLysLeuCySerThrSerLysProThrSerGluValProLeu 561  
Db 2392 TATTAAGGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 2451  
Qy 562 AlaGluAspGlyThrPhePhe 568  
Db 2452 GCGCACTTATATATTGTTTC 2472

RESULT 7  
US-08-466-248-19  
Sequence 19, Application US/08466248  
Patent No. 6258359  
GENERAL INFORMATION:  
APPLICANT: Labigne, Agnes  
APPLICANT: Sauerbaum, Sebastien  
APPLICANT: Ferrero, Richard L.  
APPLICANT: Thibierge, Jean-Michel  
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,248  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/447,177  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 435  
APPLICATION NUMBER: US 08/432,697  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495, 0137-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4400  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2619 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 31..36  
OTHER INFORMATION: /standard\_name="Shine-Dalgarno  
OTHER INFORMATION: sequence."

NAME/KEY: misc feature  
LOCATION: 756..759  
OTHER INFORMATION: /standard\_name="Shine-Dalgarno  
OTHER INFORMATION: sequence."  
US-08-466-248-19

Alignment Scores:  
Pred. No.: 5,14e-241 Length: 2619  
Score: 2244.00 Matches: 411  
Percent Similarity: 85.01% Conservative: 71  
Best Local Similarity: 72.49% Mismatches: 85  
Query Match: 74.82% Indels: 0  
Gaps: 3

US-09-904-994b-3 (1-568) x US-08-466-248-19 (1-2619)

Qy 2 LysMetLysLysGlnGluLysThrValAsnThrTyGlyProThrLysGlyAspLysValArg 21  
Db 772 AAGATTTCACGAAAGAAATATGTTCTATGTATGTGTCCTACCGGGGATCGTTTAGA 831  
Qy 22 LeuGlyAspThrAspLeuThrPalaGluValGluHisAspLysThrThrTyGlyGluGlu 41  
Db 832 CTGGCGACACTGATTTGATCTTGAAGTGAGCATGATGACACCTTATGTGAAGAG 891  
Qy 42 LeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerProAsp 61  
Db 892 ATCAATTTGGGGGGCGGTAAACCTATCCGATGATGAGTCAACCAATAGCCCTTAC 951  
Qy 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspLysThrGlyIleTy 81  
Db 952 TCTTATGAATTAAGATTGTGCTCCTACATGACCTCTCATTTGAGACTATACGGGATTTAC 1011  
Qy 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysValAsnLys 101  
Db 1012 AAAGCCGACATTTGGGATTTAAAGCGCAAGATTCGACGATTCGCAAGCGCAATTAAG 1071  
Qy 102 AspMetGlnAspGlyValSerProHisMetValGlyValGlyThrGluAlaLeuAla 121  
Db 1072 GACATGCAAGATGCGCTGATTAATATATCTTGGCTAGTCTCTGCTACAGAGCTTTGGCA 1131  
Qy 122 GlyGluGlyMetIleIleThrAlaGlyIleLysAspSerHisThrHisPheLeuSerPro 141  
Db 1132 GCTAGGGCTTGAATTTGATGACCGGTGGCATGATGATGATGATGATGATGATGATGAT 1191  
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValThrThrMetPheGlyGlyThrGly 161  
Db 1192 CAACAAATCCCTACTGCTTTTGGCCAGGGGTTTCAACATGATGAGAGGACACAGA 1251  
Qy 162 ProValAspGlyThrAsnAlaThrThrIleThrProGlyLysThrAsnLeuHisArgMet 181  
Db 1252 CTGCGGATGCGCAAGATGGAGACCATCATCTCCGAGCGCTTAATCTTAAAGATATG 1311  
Qy 182 LeuArgAlaAlaGluGluLysSerMetAsnValGlyPheLeuGlyLysGlyAsnSerSer 201  
Db 1312 TTGGGTGACGCGAAGATAGCCCATGATCTAGGCTTTTGGTAAAGGGAATGTGTCT 1371  
Qy 202 SerLysLysGlnLeuValGluGlnValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221  
Db 1372 TACGAACCTCTTACCGCATTCAGATTGAAGCAGGGCGATGTTGTTTAAATCCAGGA 1431  
Qy 222 AspThrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyAsp 241  
Db 1432 GACTGGGGAAGCACACTGACGCTATTCACCATGCTCTCAATGTCGCGAATATAGAT 1491  
Qy 242 ValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTyValAspAspThrLeu 261  
Db 1492 GTGCAATGGGTATCCACACGATATACCTTAACAGAGCGGGCTGTGAAGAAGACACCTA 1551  
Qy 262 AsnAlaMetAsnGlyAlaGluAlaIleHisAlaLeuTyHisIleGluGlyAlaGlyGlyHis 281  
Db 1552 GAGCGATTGCGCGCCACCATTCATTCATTCACATGAAGGGCGTGGGGGTGACAC 1611  
Qy 282 SerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuProSerSerThrThrPro 301

```
Db 1612 GCTCCAGATGTTATCAAAATGCGAGGGAATTAACTTACCCGCTCTACTAACCGG 1671
Qy 302 ThrileProtyrThrileanThrValaIaGluHileuAspMetIeuMetThrCyHis 321
Db 1672 ACCATTCCTCTTCAACCAAAACCTGAAGCCGAGCAGCATGATGTAAATGGTGCCAC 1731
Qy 322 HileuAspLysArgIleArgIuAaPheLugInPheSerGlnSerArgIleArgProGly 341
Db 1732 CACTTGATTAAGTATCAAGAAAGTGTGCACTTTGCCGATTCCAGAGTTGCCCCCAA 1791
Qy 342 SerIleAlaIaGluAaPvalIleuHileuAspMetGlyValIleAlaMetThrSerSerAap 361
Db 1792 ACTATGCGCGCTGAAGACCAACCATCCATGACATGGGATCTTTCTATCAACAGCTCCGAC 1851
Qy 362 SerGlnAlaMetGlyAArgIaGlyValIleProArgThrTrpGlnThrAlaAspLys 381
Db 1852 TCTCAGGCTATGGAGCGCTAGCGGAGTGATCAACGACCTTGGCAGACAGACACAA 1911
Qy 382 AaLysLysGluPheGlyLysLysLeuProGluAaPglyLysAaPaaAspAaPheArgIle 401
Db 1912 AACAAAAAGAGTTTGGCGCTTGAAGAGAAAAAGCGATACACAACTTCCGCATC 1971
Qy 402 LysArgIyrIleSerLysTrThrIleAsnProAlaLeuThrHisGlyValSerGlyTr 421
Db 1972 AACCGTACATCTCTAAATACACATCAACCCCGGATCGCGCATGGATTTCTGACTAT 2031
Qy 422 IlegLysTrValGluGluGlyLysIleAlaAspLeuValValTrpAsnProAlaPhePhe 441
Db 2032 GTGGGCTGTGTGAATGGGCAAAATACGCCGACTGTGTCTGTGGAGTCCGGCTTCTT 2091
Qy 442 GlyValLysProLysIleValIleLysGlyGlyMetValValPheSerGluMetGlyAaP 461
Db 2092 GGCATTAAAGCCCAATATGATTTAAAGGCGGATTTATTCGCTCTCTCAATAGGCGCAT 2151
Qy 462 SerAsnAlaSerValProThrProGlnProValTrpTrpArgIuMetPheGlyHis 481
Db 2152 GCCAATGCGCTATTCGCCACCCCTCAGCCGCTATTAACGTAAATGTGGACACCAT 2211
Qy 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTrpGluAaGly 501
Db 2212 GGGAAAAACAAATTCACACACATATCATCTTGTCTCCACAGCGCTTACAGAGGAGG 2271
Qy 502 ValLysGluLysLeuGlyLeuGluAArgGluValIleProValLysAsnCyAspAsnIle 521
Db 2272 ATCAAGAAAGAACTTAAGGCTAAGTCCGCGGACCCCGCACTGAATACTGCGCATATC 2331
Qy 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541
Db 2332 ACTAAAAAGACCTCAAAATTCACAGATGTGACCGCACATATGTATCAACCTGAAC 2391
Qy 542 PheGluValPheValAspGlyLysLeuCyThrSerLysProThrSerGlnValProLeu 561
Db 2392 TATTAAGTGAAGATGGCAAGAGGTAACCTTAAGCAGACAGATGAATTGAGCTTA 2451
Qy 562 AlaGlnArgTrpThrPhePhe 568
Db 2452 GCGCAACTTATATTTGTTTC 2472

RESULT 8
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Friehtauf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
```

```
STATE: New York
COUNTRY: USA
ZIP: 10016-2088
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.72mb
COMPUTER: IBM PC compatible (NEC PC-9801 RX)
OPERATING SYSTEM: MS DOS
SOFTWARE: ASCII Form
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/732,242C
FILING DATE: 19910718
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JPN 2-210178
FILING DATE: 10-AUG-1990
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Herbert
REGISTRATION NUMBER: 17081
REFERENCE/DOCKET NUMBER: 910532/HG
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 972-1400
TELEFAX: (212) 370-1622
TELEX: 236268
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 6131 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-07-732-242C-8

Alignment Scores:
Pred. No.: 3 01e-216 Length: 6131
Score: 2028.50 Matches: 375
Percent Similarity: 79.72% Conservative: 77
Best Local Similarity: 66.14% Mismatches: 112
Query Match: 67.64% Indels: 3
DB: 1 Gaps: 2

US-09-904-994B-3 (1-568) x US-07-732-242C-8 (1-6131)
Qy 3 MethylLysGluGluLysThrIleArgGluGlyMetGlyGlnSerAsn---SerProAap 22
Db 989 ATGTCCTCAAGCAATATGCGGATATGTTGGACCAACTGTGCGGACCCATTCGTTG 1048
Qy 23 GlyAspThrAspLeuTrpAlaGluValGluHisAspTrpThrThrTrpGlyGluGlu 42
Db 1049 GCAGATTTCAGAAATGTTTATCGAAATTTGAAGACTATACACGTATGAGATGAGTA 1108
Qy 43 LysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsn---SerProAap 61
Db 1109 AAGTTTGGCGGCGGAGAGTGTCCGAGTGAATGGGACAGATCCTTTGGCGACAGC 1168
Qy 62 GluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTrpThrGlyIleTr 81
Db 1169 GATGAATGCGTGCATCTGTATTAACAATGCAATATGTTGATTACACAGATTTAT 1228
Qy 82 LysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsnLys 101
Db 1229 AAAGCAGATATCGCATTAATAAAGATGATGATTCCTCATGTGAAAAAGCGGGGAACCG 1288
Qy 102 AspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGluAlaLeuAla 121
Db 1289 TTGTTATGAGCGGGGTGAT-----ATGGATTTGGAGCGACACAGAAATCATAGCC 1342
Qy 122 GlyGluGlyMetIleIleThrAlaGlyIleAspSerHisThrHisPheLeuSerPro 141
Db 1343 GCAGAAAGGATATGTTGTACACCCGAGATATGATGTCATATTCATCTTTATTTGCCCT 1402
Qy 142 GlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPheGlyGlyGlyThrGly 161
Db 1403 CAGCAATCGAAACCGCTCTTCATCGGGTGTGACCACTATGATTTGGCGAGGAAACAGA 1462
```

QY 162 ProValAspGlyThrAsnAlaThrThrIleThrProGluLysIleTyrAsnLeuHisArgMet 181  
 Db 1463 CCGCCCTACAGGCAACAAATCCACTACTGTACACGGGCGCTTGGAATATCCATCGATATG 1522  
 QY 182 LeuArgAlaAlaGluGluTyrSerMetAsnValGlyPheLeuGlyLysGlyAsnSer 201  
 Db 1523 CTTCACAGCGCCGAAGAATCCCGATTAACCTTGCGCTTTTAGGAAAGGAACCTGTCCA 1582  
 QY 202 SerLysLysGluLeuValGluGluValGluAlaGlyAlaIleGlyPheLysLeuHisGlu 221  
 Db 1583 GATAGGCTCTTTAAAGAACAAATTGAAACGGGAGCGGTGAGTTTAAACCTTCACCA 1642  
 QY 222 AspTyrGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGluTyrAsp 241  
 Db 1643 GATTGGGAGATGACGCGCGCGCTATTGATCATGTTTGAAGTGGCGGATCATATGAT 1702  
 QY 242 ValGluValCysIleHisThrAspThrValAsnGluAlaGlyTyrValAspAspThrLeu 261  
 Db 1703 GTGCAAGTAGCGATTTCATACAGACACTTAATAGAGCGGATTTGTCAGAGATACCTTG 1762  
 QY 262 AsnAlaMetLeuGlyArgAlaIleHisAlaTyrHisIleGluGlyAlaGlyGlyHis 281  
 Db 1763 AAAGCCCTAGACGCTCAGATGATTCATACCTATACAGAAAGGCGCTGCGGCGGACAT 1822  
 QY 282 SerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuProSerSerThrPro 301  
 Db 1823 GCTCCGATATTATAAAAGCGCGCTTCCGAAATTTTGCTTCTTCCACGAATCCA 1882  
 QY 302 ThrIleProTyrThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHis 321  
 Db 1883 ACTGACCTTATATCTATCAATACCTTTGGAGAGACATTTAGATGTTTATGTTTGCAC 1942  
 QY 322 HisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSerArgIleArgProGly 341  
 Db 1943 CACTACAGCCTTAATATCTCCAGAGATATTCCTTTGCGCATTCACGATACGGAAGAG 2002  
 QY 342 SerIleAlaIleGluAspValLeuHisAspMetCysValIleAlaMetThrSerSerAsp 361  
 Db 2003 ACCATCCGCGGGAAGATGTTTACATGATTTAGCCGTTTCCACATGATTTCCATGAT 2062  
 QY 362 SerGlnAlaMetGlyArgAlaGlyGluValIleProArgThrTyrPGLnThrAlaAspLys 381  
 Db 2063 TCACAGCGCATGGGCGCAGTAGAGAACTGATCTCTACGTCGCAACCGCTGACAG 2122  
 QY 382 AsnLysLysGluPheGlyLysLeuProGluAspGlyLysAspAsnAspAsnPheArgIle 401  
 Db 2123 ATGAAAAAGCAAGAGGAAAGTTTCAAGAAAGACAATGCTGTGGAGACAACTTTCGTGTG 2182  
 QY 402 LysArgTyrIleSerLysTyrThrIleAsnProAlaLeuThrHisGlyValSerGluTyr 421  
 Db 2183 AAACGTTATATGCGCAATATACATCATATCCGGCCCATGCTCATGTGATTCGCGATTAT 2242  
 QY 422 IleGlySerValGluGluGlyLysIleAlaAspLeuValIleTyrAsnProAlaPhePhe 441  
 Db 2243 GTGGGTTCTGTGAAGTGGAAAAATAGCTGATTTAGTGGTGTGAATCCGCTTTT 2302  
 QY 442 GlyValLysProLysIleValIleLysGlyCysMetValValPheSerGluMetGlyAsp 461  
 Db 2303 GGTTGTGAACCTGAACGTGCTTAAAGAGAGATGATTGCTTACAGACATAGGAGAT 2362  
 QY 462 SerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPheGlyHisHis 481  
 Db 2363 CCCATATCCAGCATTCGACACCGCACCGGTTTAAATCGTCGATGTTTGCAGCGAAA 2422  
 QY 482 GlyLysAlaLysPheAspThrSerIleThrPheValSerLysValAlaTyrGluAsnGly 501  
 Db 2423 GGAGATCCAAATATCAACGCTATACCTTTTTCGAAAGACGCTATGAAAAAGGC 2482  
 QY 502 ValLysGluLysLeuGlyLysGluArgGluValLeuProValLysAsnCysArgAsnIle 521  
 Db 2483 ATTCATGAACAGTTGGTTGAAGAAAAAGGTGAACACAGTCATGGAATTCGAAAAATTG 2542

QY 522 ThrLysLysAspPheLysPheAsnAspLysThrAlaLysIleThrValAspProLysThr 541  
 Db 2543 ACGAAAAAGATTTTAATTTTGAACGATTAACCCCAAAATTTAGAGTCGATCCACAGCA 2602  
 QY 542 PheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGlnValProLeu 561  
 Db 2603 TATGAATGAAGAGTAGACGCGTCATATGATGACATGTGAACCGGAGAAAATGTCCTATG 2662  
 QY 562 AlaGlnArgTyrThrPhePhe 568  
 Db 2663 GCACACGCTAATTTCTTATTT 2683  
 RESULT 9  
 US-09-252-991A-10316  
 ; Sequence 10316, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; PRIOR FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 10316  
 ; LENGTH: 1701  
 ; TYPE: DNA  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-10316  
 Alignment Scores:  
 Pred. No.: 1,16e-194 Length: 1701  
 Score: 1828.00 Matches: 336  
 Percent Similarity: 74.47% Conservative: 87  
 Best Local Similarity: 59.15% Mismatches: 143  
 Query Match: 60.95% Indels: 2  
 Gaps: 1  
 US-09-904-994b-3 (1-568) x US-09-252-991A-10316 (1-1701)  
 QY 1 MetLysMetLysLysGluGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20  
 Db 1 ATGAAAAATCGACGACAAAGCGTACGCCGACATGTTCCGCCCCACCGCGCGGAGCGGTG 60  
 QY 21 ArgLeuGluAspThrAspLeuThrAlaGluValGluHisAspTyrThrTyrGlyGlu 40  
 Db 61 CGCTGGCCGACACCGCATCGTGTGATCGAGGTGGAAACGGGACTTCACCGCTATGGCGAG 120  
 QY 41 GluLeuAspPheGlyAlaGlyLysThrIleArgGluGlyMetGlyGlnSerAsnSerPro 60  
 Db 121 GAAGTGAAGTTCCGCGCGCGCAAGTCATCCGACACGCGATGGCGCCAGATCAGCTGGCC 180  
 QY 121 GAAGTGAAGTTCCGCGCGCGCAAGTCATCCGACACGCGATGGCGCCAGATCAGCTGGCC 180  
 Db 61 AspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAspTyrThrGlyIle 80  
 Db 181 GCGCGCAGAGTGGCGACAGGTGATGACCAATGCGCTGATCCCTCGACCACTGGGCGGTG 240  
 QY 81 TyrLysAlaAspIleGlyLysLysValAsnGlyLysIleHisGlyIleGlyLysAlaGlyAsn 100  
 Db 241 GTCAAGGCCGACGTCGCTCTCAAGAGAGGGGCAATCCAGGCCATCCGCAAGGCGGCAAC 300  
 QY 101 LysAspMetCysAspGlyValSerProHisMetValValGlyValGlyThrGluAlaLeu 120  
 Db 301 CCGTGAATCCAGCCCGCGGTG-----ACATCGCCCATCGCGCGCGGACGAGTATC 354  
 QY 121 AlaGlyGluGlyMetIleIleThrAlaGlyGlyIleAspSerHisThrHisPheLeuSer 140  
 Db 355 GCGCGCAGAGGAGTGAATCCCTCACCGCGCGGATCGACACGCAATCCACTCATCTGC 414  
 QY 141 ProGlnIleProThrAlaLeuAlaAsnGlyValIleThrThrMetPheGlyGlyThr 160

Db	415	CGCGACGATCGAAGAGCGCTGATGAGCGGGGTCACACCATGATCGGGGGCACCC	474
Qy	161	GIYProValaAergLYThrAsnAlaThrThrILeThrProGLYsTRPAsnLeuHiSaArg	180
Db	475	GGCCCCGGCGCGGGAGCCAGCCACGACCTGCACTCGGGTCCCTGGCACATGGCGGG	534
Qy	181	MetLeuAArgAlaAlaGLuGLuTYrSerMetAsnValGlyPheLeuGLYVgLYAsnSer	200
Db	535	ATGCTCCAGCGCGCGCGAAGCCCTCCGATGAACTCGGCTTCACCGGCAAGGCAACCC	594
Qy	201	SerSerLYsLYsGLuInLeuValGLuGLuValaGLuAlaGLYAlaGLYPhaLYLeuHiS	220
Db	595	AGCCTGCGCGCTCGCGCTGAGAGAGAGGTGGCTCGCGGCGCATCGGCTTAAGCTGAC	654
Qy	221	GLuAspTRPGLYThrThrProSerAlaILeAspHiScYsLeuSerValAlaAspGLYTr	240
Db	655	GAGAGACTGGGCGAGACACCCGGCGCGCATTCGACAACTCCCTGAGAGTGGCGAGCGCAC	714
Qy	241	AspValaGLuValCYsVILeISThrAspThrValaAsnGLuAlaGLYTYrValaAspAspThr	260
Db	715	GACATCCAGTGGCGATCCACACGACCCCTCAACGAGTCCGGCTTGTCGACGACCC	774
Qy	261	LeuAsnAlaMetAsnGLYArgAlaILeHisAlaTYrHiSILeGLuGLYValaGLYGLY	280
Db	775	CTCGGCGCTTCAAAGGGCGGACCATTCACACTTACACACCGAGAGGGCGCGGGCGCG	834
Qy	281	HisSerProAspValILeThrMetAlaGLYLeuAsnILeLeuProSerSerThrThr	300
Db	835	CACGCACCGGACATCATCAAGGCCCGCGCTTCGCCAACGTGTCCAGCTCGACCAAC	894
Qy	301	ProThrILeProTYrThrILeAsnThrValAlaGLuHISLeuAspMetLeuMetThrCYs	320
Db	895	CCGACACCGCGCGTTCACCCGCAATPACATCCAGCAGCACCTGACATCTGATGGTGC	954
Qy	321	HisHISLeuAspLYsArgILeArgGLuAspLeuGLuInPheSerGLuSerArgILeArgPro	340
Db	955	CACCACTCGACCGCGCATCGCGGACGAGAGCGTGGCTTGCGGATCGCGATCGCGCGC	1014
Qy	341	GLYSerILeAlaAlaGLuAspValaLeuHiSaAspMetGLYAlaILeAlaMetThrSerSer	360
Db	1015	GAGACCATCGCGCGAGACATCTCCACGACCTCGGCGGGTTCAGATATACGCTCC	1074
Qy	361	AspSerGLuAlaMetGLYArgAlaGLYGLuValILeProArgThrTRPGLuThrAlaAsp	380
Db	1075	GACAGCCAGCGCATCGCGGGGTGGCGAAGTATCAGCGCCACTCGGCACACCGCCGAC	1134
Qy	381	LYsAsnLYsLYsGLuPheGLYLYsLeuProGLuAspGLYLYsAspAsnAspAsnPheArg	400
Db	1135	AAGATGAAGGGCCAGCGCGCGCTTCACCGGCGGATGGCGGACACCACTTCGCG	1194
Qy	401	ILeLYsArgTYrILeSerLYsTYrThrILeAsnProAlaLeuThrHiSGLYValaSerGLu	420
Db	1195	GCCAGCGGCTCATGCGCCAAAGTACCATCAACCGCGGATCAACCCAGCATACGCAT	1254
Qy	421	TYrILeGLYSerValaGLuGLuGLYLYsILeAlaAspLeuValaValaTRPAsnProAlaPhe	440
Db	1255	GAAAGGGGCTCCGTGGAAGCGGCAAGGGGCGGACCTGTGGTGGCGCGCGCTTC	1314
Qy	441	PheGLYValaLYsProLYsILeValILeLYsGLYGLYMetValaILePheSerGLuMetGLY	460
Db	1315	TTTCGCGTCGAAGCGAGCTGATCTCTCAAGGGCGGGGACATCGCGCAGCTGATGGGC	1374
Qy	461	AspSerAsnAlaSerValaProThrProGLuProValaTYrTYrArgGLuMetPheGLYHiS	480
Db	1375	GACATCAACGGCTGATCTCCACGCGCGAGCGGGTGTGCATTCACCAAGCATTTGCCACGC	1434
Qy	481	HisGLYLYsAlaLYsPheAspThrSerILeThrPheValaSerLYsValaAlaTYrGLuAsn	500
Db	1435	TACGCGGCGAGCGGCACGCGACGAGCTTGATCGTCAAGGAGCGGCTTGCGCGCC	1494
Qy	501	GLYValaLYsGLuLYsLeuGLYLYsLeuGLuArgGLuValaLeuProValaLYsAsnCYeArgAsn	520

Db	1495	GGCGTACCGCAGCAGCTCGCGCTCGCAGAGGCCATCGCGCGTGGTACGCGCTCGCGGT	1555
Qy	521	11ethrlylsyaaaphehlyspheahsapslythralalysliethrValaspProlys	540
Db	1555	GTGCGAAGAACCGCACTGTATCCACAACGGCTACTACCGACCATCGAGTGGACGACAG	1614
Qy	541	ThrhhegluValpheValaspGlylybheucysrthrserysProthrsesgluValPro	560
Db	1615	AACRACCAAGTACCGCGCCGAGCGGACAGTGGCTCTGTGTCGGAACCGGCGGACGTGTCGG	1674
Qy	561	leuAlaInArGlyTyrThrPhePhe	568
Db	1675	ATGGCGCAGCGCTATTTCCTCTTC	1698
RESULT 10			
	US-08-487-429A-3		
	/ Sequence 3, Application US/08487429A		
	/ Patent No. 6468765		
	/ GENERAL INFORMATION:		
	/ APPLICANT: Fleischmann et al.		
	/ TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus influenzae Rd		
	/ Patent No. 6468765		
	/ TITLE OF INVENTION: Genome, Fragments Thereof, and Uses Thereof		
	/ FILE REFERENCE: PB186P1		
	/ CURRENT APPLICATION NUMBER: US/08/487,429A		
	/ PRIOR FILING DATE: 1995-06-07		
	/ PRIOR APPLICATION NUMBER: 08/426,787		
	/ PRIOR FILING DATE: 1995-04-21		
	/ NUMBER OF SEQ ID NOS: 10		
	/ SOFTWARE: PatentIn Ver. 2.1		
	/ SEQ ID NO 3		
	/ LENGTH: 1716		
	/ TYPE: DNA		
	/ ORGANISM: Haemophilus influenzae		
	US-08-487-429A-3		
	Alignment Scores:		
	Pred. No.: 1,97e-194	Length: 1716	
	Score: 1826.00	Matches: 337	
	Percent Similarity: 75.17%	Conservative: 93	
	Best Local Similarity: 58.92%	Mismatches: 136	
	Query Match: 60.89%	Indels: 6	
	DB: 3	Gaps: 3	
	US-09-904-994B-3 (1-568) x US-08-487-429A-3 (1-1716)		
Qy	1	MethylmethylslyslngluTyrValasnThrTyrGlyProthrlYgIyAspIysVal	20
Db	7	TTAACAAATTTCAGAGCGCCAAATATGTACCACTTATGCTCCAAAGTGGCGATTAAGTC	66
Qy	21	ArgleugIyAspThrAspLeuTrrAlaGluValGlnHisAspTyrThrThrTyrGlyGlu	40
Db	67	CGTTTAGCGCATCAACAAATTATGGGCACCACTTGAACAAAGATTATATGACCAAGAGTGAT	126
Qy	41	GluLeuIysPheGlyAlaGlyLyserThrIleArgGluGlyMetGlyGlnSerAsnSer---	59
Db	127	GAGGTAAATTATGGTGGCGGTAAAGCGTGGTGTATGGCTCAAGCGGTACGGCA	186
Qy	60	-----ProAspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleIleAsp---	76
Db	187	ACTGGCGCAACATCCGAATGTATTTGGATTTTGTGATTACCAACGATATCATTTGATGCT	246
Qy	77	TyrThrGlyIleTyrLyysAlaAspIleGlyIleLyAsnGlyLyylleHleGlyTleGly	96
Db	247	AAATTAGGCATTTCGAAGCGCATTTGGTATTTGGATGGGCGGATTTGGGTATTTGSA	306
Qy	97	LyysAlaGlyAsnLyAspMetGlnAspGlyValserProHleMetValAlaGlyValGly	116
Db	307	CAAGCAGGTAAACCCGTACACCATGATTAACGTACACCAACAAATATGATTTATCGTCAAGC	366
Qy	117	ThrGluAlaLeuAlaGlyGluGlyMetIleIleThrAlaGlyGlyTleAspSerHisThr	136
Db	367	ACGGAAAGTTCATTAACGGTGTCACATTATATGCACACCGCTGGTGGTATGATCCACATTT	426

```

Qy 137 HisPheLeuSerProGlnGlnPheProThrAlaLeuAlaGlnGlyValIleThrThMetPhe 156
Db 427 CACTTTATTTTGTCCACACACACACACACATCATCAATGAAAGTGGGTATACACGTTAAATT 486
Qy 157 GlyGlyGlyThrGlyProValAspGlyThrAsnAlaThrThrIleThrProGlyLysTrp 176
Db 487 GGTGGTGAAGTGGCCCTGGTGGTATGATACACACGCAACCATTTGATACCCCTGGGCGATGG 546
Qy 177 AsnLeuHisArgMetLeuArgAlaAlaGluGlyTrpSerMetAsnValGlyPheLeuGly 196
Db 547 TATATGAAAGCTAATGTTTCAAGCCGACAGACCTTGGCGGTAAACGTCGGATTTTGTGGT 606
Qy 197 LysGlyAsnSerSerSerLysGlyGlnLeuValGluGlnValAlaGluAlaIleGly 216
Db 607 AAAGGCAACTGTTTCAACCCCTTATGATCCCTGGTGAATGAAAGCGGATGATTAAGT 666
Qy 217 PheLeuLeuHisGlyAspTrpGlyThrThrProSerAlaIleAspHisCysLeuSerVal 236
Db 667 TTAATAATCCACGAAAGATGGGGTGCAACGCCCTGGATTTGATTTGCTTAAAGTA 726
Qy 237 AlaAspGlyTrpAspValGlnValCysIleHisThrAspThrValAsnGluAlaGlyTrp 256
Db 727 GCAAGTAAATGGATATTCAGTGGCCATTCACACAGACGCTAAATGAAAGTGGCTTT 786
Qy 257 ValAspAspThrLeuAsnAlaMetAsnGlyValArgAlaIleHisAlaTrpHisIleGluGly 276
Db 787 TTGGAAAGACGAGTAAAGCATTTGATGAGACGAGTCACTTCACTTCTTCAACGAGGAGC 846
Qy 277 AlaGlyGlyGlyHisSerProAspValIleThrMetAlaGlyGluLeuAsnIleLeuPro 296
Db 847 GCAAGTGGTGGTATGACCATTCATCAATTAAAGCAGATGATTAACAGATTAACCT 906
Qy 297 SerSerThrThrProThrIleProTrpThrIleAsnThrValAlaGluHisLeuAspMet 316
Db 907 GCTTCAACCAACCCGACTCGCTCCCTTTACCAAAAACCATTTGATGAAATTTGATGATG 966
Qy 317 LeuMetThrCysHisHisIleuAspLysArgIleArgIleuAspLeuGlnPheSerGlnSer 336
Db 967 TTGATGGTTTGGCCATCTTATGATTAACGCTGGGAAAGCGTATTTGGCGATGAC 1026
Qy 337 ArgIleArgProGlySerIleAlaAlaGluAspValLeuHisAspMetGlyValIleAla 356
Db 1027 CGTATCGCCCTGAAACATTTGACGAGAGATATTTTGCATGATATCGGCTCTTCTCC 1086
Qy 357 MetThrSerSerAspSerGlnAlaMetGlyValArgAlaGlyGluValIleProArgThrTrp 376
Db 1087 ATTATGAGTTCAGACTCTCAAGCGATGGGACGATTTGGCGAAAGTCGTTATTCGTAACATGG 1146
Qy 377 GlnThrIleAspLysAsnLysGlyGluPheGlyLysLeuProGluAspGlyLysAspAsn 396
Db 1147 CAATCTGCAATTAAGTAAATGCAACGTGGTATGATTAAGGTAAGGA-----AAC 1200
Qy 397 AspAsnPheArgIleLysArgTrpIleSerLysTrpThrIleAsnProAlaLeuThrHis 416
Db 1201 GATTAACCTCGTATTAACGATATATCGCAATATACCAATCAACCCAGCATTTGCACAT 1260
Qy 417 GlyValSerGlyTrpIleGlySerValGluGluGlyLysIleAlaAspLeuValValTrp 436
Db 1261 GGTAATTCGAGACATATTTGGCTCGTTAGAAAGTGGTAAATCGACAGATATCGTGTATGG 1320
Qy 437 AsnProAlaPhePheGlyValLysProLysAlaIleValIleLysGlyGlyMetValValPhe 456
Db 1321 AAACCGATGTTCTTTGGCGTAAACCTGAAGTGGTATTAATAAAGGCTTATTAAGCAT 1380
Qy 457 SerGluMetGlyLysSerAsnAlaSerValProThrProGlnProValTrpTrpArgGlu 476
Db 1381 GCGAAATGGCGCATTCGAATGCTCTCAATTCACACCGCACTGATTTTACCGTCCA 1440
Qy 477 MetPheGlyHisHisGlyLysAlaLysPheAspThrSerIleThrPheValSerLysVal 496
Db 1441 ATGTACGTCGACAAAGGCTTGAACACGCAACAAACAGCAGATATCTTTGTTCAACAGCC 1500

```

```

Qy 497 AlaTrpGluAsnGlyValLysGluLysLeuGlyLeuGluLubArgGlnValLeuProValLys 516
Db 1501 GCTGAAAAGCTGATATATTCGGCGAAGTTCGTTTACCAAAAGAACCATTTGCTGAAA 1560
Qy 517 AsnCysArgAsnIleThrLysLysAspPheAsnAspLysThrAlaLysIleThr 536
Db 1561 GGTCGCCGCAACGATGATTAAGTAAAGATCTGGTTCATATGATGTAACACCAACATTAAT 1620
Qy 537 ValAspProLysThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThr 556
Db 1621 GTTATCTCTGAACCTTATGAAAGTTGAGTGGACGAGAGTTAATTAACCTGTGAACAGTG 1680
Qy 557 SerGlnValProLeuAlaGlnArgTrpThrPhePhe 568
Db 1681 GATAGCGTACCATTTGGGTACGCAATATTTCTTATTC 1716

RESULT 11
PCT-US96-05320A-541
Sequence 541, Application PC/TUS9605320A
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences
APPLICANT: 9410 Key West Avenue
APPLICANT: Rockville, MD 20850
APPLICANT: United States of America
APPLICANT: Johns Hopkins University
APPLICANT: 720 Rutland Avenue
APPLICANT: Baltimore, MD 21205
APPLICANT: United States of America
APPLICANT: Mark D. Adams
APPLICANT: Owen White
APPLICANT: Hamilton O. Smith
APPLICANT: J. Craig Venter
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome,
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20003-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05320A
FILING DATE: April 22, 1996
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/476,102
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,429
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Eric K. Steffe
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488.014PC01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 541:
SEQUENCE CHARACTERISTICS:
LENGTH: 1716 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
PCT-US96-05320A-541

Alignment Scores: 1.97e-194 Length: 1716
Pred. No.:

```



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/557,884  
FILING DATE: 25-Apr-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/476,102  
FILING DATE: JUN-5-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Michelle S. Marks  
REGISTRATION NUMBER: 41,971  
REFERENCE/DOCKET NUMBER: PB186P3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-309-8504  
TELEFAX: 301-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-557-884-1

Alignment Scores:  
Pred. No.: 1,586-189 Length: 1830121  
Score: 1826.00 Matches: 337  
Percent Similarity: 75.17% Conservative: 93  
Best Local Similarity: 58.92% Mismatches: 136  
Query Match: 60.89% Gaps: 6  
Indels: 3

US-09-904-994b-3 (1-568) x US-09-557-884-1 (1-1830121)

QY 1 MetLysMetLysGlnGluTyrValAsnThrTyrGlyProThrLysGlyAspLysVal 20  
Db 564173 TTAACAATTCACAGCGCATATGTAGCACTTATGTCACAGCTGGCATTAAGTC 564114  
QY 21 ArgLysGlnAspThrAspLysLeuValGlnHisAspLysTyrThrTyrGlyGlu 40  
Db 564113 CGTTAGCGCATACCAATTATGAGCAACCATTAAGCAAGTTATGACCAAGAGTAT 564054  
QY 41 GlnLeuLysPheGlyValGlyLysThrLeuArgGlnLysMetGlyLysSerAsnSer--- 59  
Db 564053 GAGGTAAATTTGGTGGCGGAAAGCGTGGTATGCTGCTCAAGCGCGTAAAGCA 563994  
QY 60 -----ProAspGlnAsnThrLeuAspLeuValLeuThrAsnAlaMetLeuLeuAsp--- 76  
Db 563993 ACTGCGACAAATCGAATGTATGATTTGTGATTACCAAGTATGATCATGATGCT 563934  
QY 77 TyrThrGlyLeuTyrValAspLysValLeuLysGlnLysLeuValGlnValGly 96  
Db 563933 AAATTAAGCATTAACCAAGCGATATGATGATGATGATGATGATGATGATGATGAT 563874  
QY 97 LysAlaGlnLysAsnLysAspMetGlnAspGlyValSerProHisMetValGlnValGly 116  
Db 563873 CAAGCAGTAACTCTGACACCATGATACCTCAACCAATATGATATGATGATGATGAT 563814  
QY 117 ThrGlnAlaLeuAlaGlnLysGlnLysMetLeuLeuThrAlaGlnLysLysSerHisThr 136  
Db 563813 ACGGAACTTCAACGCGTGCATTTAATTCACACCGTGGTATGATGATGATGATGAT 563754  
QY 137 HisPheLeuSerProGlnGlnPheProThrAlaLeuAlaGlnLysValThrThrmecPhe 156  
Db 563753 CACTTATTTGTCACCAACCAAGCAACCATGCAATGAAAGTGGCGTTTACACCGTTAAT 563694  
QY 157 GlyGlyGlyThrGlyProValAspGlyThrAsnAlaThrThrLeuProGlyLysThr 176  
Db 563693 GGTGGTGAAGTGGCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 563634  
QY 177 AsnLeuHisArgMetLeuArgAlaAlaGlnLysTyrSerMetAsnValGlyPheLeuGly 196  
Db 563633 TATATGAACGATATGTTCAAGCGGCAAGACCTTGGCGTAAACGTCGATTTTGGT 563574

QY 197 LysGlyAsnSerSerSerLysGlnLeuValGlnGlnValGlnAlaGlyValAlaGly 216  
Db 563573 AAAGCAGTCTTACACCTTATGATCTCTGGTGAAGCAATTTGAAGCGGTCATGAT 563514  
QY 217 PheLysLeuHisGlnAspTyrGlyThrThrProSerAlaLeuAspHisCysLeuSerVal 236  
Db 563513 TTAATAATTCACAGAACTGGGGGTCAGAACCGCTCGGATGATGATGATGATGATGAT 563454  
QY 237 AlaAspGlyTyrAspValGlnValCysLeuHisThrAspThrValAsnGlnAlaGlyTyr 256  
Db 563453 GCACATGAATGATGATTCACAGTGGCCATTCACACAGACGCTAAATGAAAGCGCTTT 563394  
QY 257 ValAspAspThrLeuAsnAlaMetAsnGlyArgAlaLeuHisAlaTyrHisAlaGlnGly 276  
Db 563393 TTGAAGACACAGATGAAGAAAGCATTTGATGACAGATCTTCATCTTCATTCAGAGGCG 563334  
QY 277 AlaGlyGlyGlnHisSerProAspValLeuThrAlaGlnLysLeuAsnLeuPro 296  
Db 563333 GCAGTGGTGGTCAATGACCTGACATCATTAAGCAGATGATTCACACGATTTAAGCT 563274  
QY 297 SerSerThrThrProThrLeuProTyrThrThrLeuAsnThrValAlaGlnHisLeuAspMet 316  
Db 563273 GCTTCAACCAACCGACTCGCTCGCTTACCAAAACACATGATGAAACATTTGGATATG 563214  
QY 317 LeuMetThrCysHisLeuAspLysArgLeuArgGlnAspLeuGlnPheSerGlnSer 336  
Db 563213 TTGATGGTTGGCAATCACTTATGATTAACGGTGGCGGAAGCGTATGCTTTGGCATAGC 563154  
QY 337 ArgLeuArgProGlySerLeuAlaAlaGlnAspValLeuHisAspMetGlyValAlaAla 356  
Db 563153 CGTATCCGCGCTGAACCATTTGACGAGAAAGATATTTTGGATGATATGGGCGCTTCC 563094  
QY 357 MetThrSerSerAspSerGlnAlaMetGlyArgAlaGlnGlyValLeuProArgThrTrp 376  
Db 563093 ATTATGATGTTGACAGCTCTCAAGCATGAGAGCTATTTGGCGAAGCTGATATTCGACATG 563034  
QY 377 GlnThrAlaAspLysAsnLysGlnPheGlyLysLeuProGlnLysAspAsp 396  
Db 563033 CAACCTGACATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 562980  
QY 397 AspAsnPheArgLeuLysArgTyrLeuSerLysTyrThrLeuAsnProAlaLeuThrHis 416  
Db 562979 GATTAATCTTCGATTAACGATATATCGCAATATACCATCAACCGACATTTGCAT 562920  
QY 417 GlyValSerGlnTyrLeuGlySerValGlnGlnLysLysLeuAlaAspLeuValTyr 436  
Db 562919 GGTATTCGAGGACATTTGGCTCGTGAAGTGGGTAAATTCGACATATCGTGTATGG 562860  
QY 437 AsnProAlaPhePheGlyValLysProLysAlaLeuLeuGlyGlyMetValValPhe 456  
Db 562859 AAACCGATGTTCTTGGCGTAAACCTGAAGTGTGATTAATAAGGCTTTATTAAGCTAT 562800  
QY 457 SerGlnMetGlyAspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGln 476  
Db 562799 GCGAAATGGCGCATCCAAATGCTCAATTCACACCGACCTGATTTACCGCTCA 562740  
QY 477 MetPheGlyHisGlnGlyLysValAlaLysPheAspThrSerLeuThrPheValSerVal 496  
Db 562739 ATGACAGCTGACAAAGCTTATGACAAACGCAACAGATATCTTTGTTTACACAGCC 562680  
QY 497 AlaTyrGlnLysGlnValLysGlnLysLeuGlyLysLeuGlnGlnValLeuProValLys 516  
Db 562679 GCTGAATAACCTGATATTCGTCGCAACATTCGTTTACAAAGAAACCATTTGCTGTA 562620  
QY 517 AsnCysArgAsnLeuThrLysLysAspPheLysAsnAspLysThrAlaLysLeuThr 536  
Db 562619 GCGTGGCGGCAACGATGATTAAGATCTGGTTCATATGATGTAACCAACCAATTA 562560  
QY 537 ValAspProLysThrPheGlnValPheValAspGlyLysLeuCysThrSerLysProThr 556  
Db 562559 GTTATCTGTAACCTTATGAAAGTTCGAGTGAAGAGTAAATTAATCTGTAACCAAGT 562500  
QY 557 SerGlnValProLeuAlaGlnArgTyrThrPhePhe 568

Db 562499 GATAGCGTACCATTTGGTGCAGCATATTTCTTATTC 562464  
RESULT 13  
US-09-643-990A-1/c  
Sequence 1, Application US/09643990A  
Patent No. 6528289  
GENERAL INFORMATION:  
APPLICANT: Robert D. Fleischmann  
Mark D. Adams  
Owen White  
Hamilton O. Smith  
J. Craig Venter  
TITLE OF INVENTION: The Nucleotide sequence of  
the Haemophilus influenzae Rd Genome, Fragments  
Thereof, and Uses Thereof  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville,  
STATE: MD  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3 1/2 inch diskette  
COMPUTER: Dell Pentium  
OPERATING SYSTEM: MS DOS v6.22  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/643,990A  
FILING DATE: 23-Aug-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/487,429  
FILING DATE: 1995-06-07  
APPLICATION NUMBER: 08/426,787  
FILING DATE: 1995-04-21  
ATTORNEY/AGENT INFORMATION:  
NAME: Kenley K. Hoover  
REGISTRATION NUMBER: 40,302  
REFERENCE/DOCKET NUMBER: PB186P1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 301-610-5790  
TELEFAX: 310-309-8439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1830121 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-643-990A-1  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
Alignment Scores:  
Pred. No.: 1,58e-189 Length: 1830121  
Score: 1826.00 Matches: 337  
Percent Similarity: 75.17% Conservative: 93  
Best Local Similarity: 58.92% Mismatches: 136  
Query Match: 60.89% Indels: 6  
Gaps: 3  
US-09-904-994B-3 (1-568) x US-09-643-990A-1 (1-1830121)  
QY 1 MetlyMeltyllysglnGlyrValaenThrTyGlyProthrlYsglyAspLyVal 20  
Db 564173 TTAAACAATTTCAAGACGCAATGTGAGCAACTTATGTCCAACAGTTGGCGATTAAGTC 564114  
QY 21 ArgLeuGIYAAspThrAspLeuTrpAlaGluValaGluHHisAspTyrrThrThrTyrglyGlu 40  
Db 564113 CGTTTGGCGCATACCAATTTATGGGCACCACTTGAACAGATTTATTTGACCAAAAGGTGAT 564054  
QY 41 GluLeuLysPheGlyAlaGlyLysThrIleArgGluGlyMetGlyLysIserAsnSer--- 59

Db 564053 GAGTGTAAATTTGTGCGCGGTAAAGCGCGTGTATGGTTCAAAGCCGTACGGCA 563994  
QY 60 -----ProAspGluAsnThrIleuAspLeuValIleThrAsnAlaMetIleIleAsp--- 76  
Db 563993 ACTCGGACAACTCCGAACTATTTGATTTGGATTTACCAAGCTGATGATGATGCT 563934  
QY 77 TyrThnglyIleTyrLysAlaAspIleGlyIleLysAsnGlyLysIleHisGlyIleGly 96  
Db 563933 AAATTAGGATTTATGCAAGCCGATATTGGTATTCGTAATGGGCGATTTGGTATGGA 563874  
QY 97 LysAlaGlyAsnLysPheMetGlnAspGlyValIserProHisMetValaGlyValaGly 116  
Db 563873 CAAGCAGGTAAACCCGTGACACCATGTGATTAACGACCAAAATATGATTTCCGTGCAAC 563814  
QY 117 ThrGluAlaLeuAlaGlyGluGlyMetIleIleThrAlaGlyIleLysPheHisThr 136  
Db 563813 ACGGAAGTTTCATACGCGTCACATTTAATTCGACCGCTGGTGTATGATACCCACATT 563754  
QY 137 HisPheLysSerProGlnGlnPheProThrAlaLeuAlaAsnGlyValIleThrMetPhe 156  
Db 563753 CACTTATTTTGTCCACACACACACACATGCAATTTGAAGGGCGTTACACGTTAAT 563694  
QY 157 GlyIleGlyIleThrGlyProValaAspGlyThrAsnAlaThrIleThrProGlyLysTrp 176  
Db 563693 GGTGTGTGAACCTGGCCCTGCTGATGTGACACACGACACCATTTGACCCCTGGCGCATGG 563634  
QY 177 AsnLeuHisArgMetLeuArgAlaAlaGlyLysTrpSerMetAsnValaGlyPheLeuGly 196  
Db 563633 TATATGAAACGTATATTTTAAACCGGACGACGCTTGGCCGATTAACGTCGATTTTGGT 563574  
QY 197 LysGlyAsnSerSerSerLysLysGlnLeuValaGluGlnValaGlyAlaIleGly 216  
Db 563573 AAAGGAACGTGTTCACACCTTATGATCCTCGCTGAGCAAAATTTGAAGCGGTCATTAGCT 563514  
QY 217 PheLysLeuHisGlnAspTrpGlyThrTrpSerAlaIleAspHisCysLeuSerVal 236  
Db 563513 TTAAAAATCCACGAAGACTGGGGTGCACACGCTGCGTATGATTCGCTTAAAGTA 563454  
QY 237 AlaAspGluTrpAspValaGlnValaCysIleHisThrAspThrValaAsnGluAlaGlyTrp 256  
Db 563453 GCAAGATGAATGATATTTAAATGAGCCATTCACACGACACCTTAATGAAAGTGCTTT 563394  
QY 257 ValaAspAspThrLeuAsnAlaMetAsnGlyArgAlaIleHisAlaTyrHisIleGly 276  
Db 563393 TTGGAAGACAGATGAAGCGATGATGAGCAGACATTCATACCTTCATACGAGGCG 563334  
QY 277 AlaGlyGlyIleHisSerProAspValIleThrMetAlaGlyLysLeuAsnIleLeuPro 296  
Db 563333 GCAGGTGGTGTATGACACCTGACATCTTAAGACGCGATGTATTCAAACGTAATTAACCT 563274  
QY 297 SerSerThrThrProThrIleProTyrrIleAsnThrValaAlaGluHisIleuAspMet 316  
Db 563273 GCTTCAACCAACCCGACCTGCTCGTTTACCAAAACACCATGATGAACATTTGATTAAG 563214  
QY 317 LeuMetThrCysHisIleLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnSer 336  
Db 563213 TTGATGTGTTTGCATCATCTATATTAACCGCGCGGAAGACGTGTTGGCCGATAGC 563154  
QY 337 ArgIleArgProGlySerIleAlaAlaGluAspValLeuHisAspMetGlyValaIleAla 356  
Db 563153 CGTATCCGCCCTGGAACCATTCACACAGATATTTTTCATGATGATGGCCCTCTTCC 563094  
QY 357 MetThrSerSerAspSerGlnAlaMetGlyArgAlaGlyGluValaIleProArgThrTrp 376  
Db 563093 ATTATGAGTTCAGACTCTCAACCGAGGACGATTTGGGAAGTGTATTTGTCATAG 563034  
QY 377 GlnThrAlaAspLysAsnLysLysGluPheGlyLysLeuProGluAlaAspGlyLysAspAsn 396  
Db 563033 CAAATCGCAGATTAAGATGAATGACACGTGTGAGCTAGCTAAGTAAGAGA-----AAC 562980  
QY 397 AspAsnPheArgIleLysArgTyrrIleSerLysTyrrIleAsnProAlaLeuThrHis 416



Db 562979 GATTAATCCGATTAAACGATATATCGGAATAACACCATCAACCCGCAATTGCACAT 562920  
 Qy 417 GlyValSerGluTrpIleGlySerValGluGluGlyIleAlaAspLeuValTrp 436  
 Db 562919 GGTATGGGAGACATATGGCTCGTTAGAAAGGGTAAATGCGATATCGTTATAG 562860  
 Qy 437 AsnProAlaPhePheGlyValIlePheProIleValIleIleGlyIleValIleVal 456  
 Db 562859 AAACCGATGTTCTTGGCGTAAACCTGAAGTGGATTAATAAAAGGCTTATTGCTAT 562800  
 Qy 457 SerGluMetGlyAspSerAsnAlaSerValProThrProGlnProValIleTrpArgGlu 476  
 Db 562799 GCGAAATGGCGAGTCAAAATGCTCAATTCACACCGCAACCTGATTCACCGTCA 562740  
 Qy 477 MetPheGlyHisIleGlyValAlaIleAspPheAspThrSerIleThrPheValSerIleVal 496  
 Db 562739 ATGTACGCTGCACAAAGCTTATGCAACCCGCAACACAGCATTTTGTGTTTTCACAAAGCC 562680  
 Qy 497 AlaTrpGluAsnGlyValIleGlyIleGlyIleGlyIleGlyIleGlyIleGlyIleGly 516  
 Db 562679 GCTGAAAAGCTGATATTCTGTGCGAAAGTTCCGTTTACAAAGAACCATTCGTTGAAA 562620  
 Qy 517 AsnGlySerGluAsnIleThrIleGlyIleAspPheIleAsnAspIleThrAlaIleIleThr 536  
 Db 562619 GGCTGCGCGCAACGATGAATAAAAGATCTGTTCAATGATGTAACCAACCAATTTACT 562560  
 Qy 537 ValAspProIleThrPheGluValIlePheValAspGlyIleGlyIleGlyIleGlyIleGly 556  
 Db 562559 GTTATGCTGAACGTTATGAAGTTCGAGTGGACGAGAGTTAATTAATCTGTGAACCAAGTG 562500  
 Qy 557 SerGlnValProLeuAlaGlnArgTrpThrPhePhe 568  
 Db 562499 GATAGCGATCATTTGGGTCAAGCATATTTCCTATTTC 562464  
 RESULT 14  
 US-09-489-039A-2045  
 ; Sequence 2045, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; PRIORITY FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 2045  
 ; LENGTH: 1878  
 ; TYPE: DNA  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-2045  
 Alignment Scores:  
 Pred. No.: 8.3e-194 Length: 1878  
 Score: 1821.00 Matches: 338  
 Percent Similarity: 74.36% Conservative: 84  
 Best Local Similarity: 60.04% Mismatches: 139  
 Query Match: 60.72% Indels: 2  
 Gaps: 1  
 US-09-904-994B-3 (1-568) x US-09-489-039A-2045 (1-1878)  
 Qy 3 MetIleValSerGluTrpIleGlySerValGluGluGlyIleAlaAspLeuValTrp 22  
 Db 13 ATTCACGCGCGAGCTATGCGATATGTTCCGCCCAACCGCGGAGCAAGATGCGCTG 72  
 Qy 23 GlyAspThrAspLeuTrpAlaGluValGluHisAspTrpThrTrpIleGlyIleGluLeu 42  
 Db 73 GCGGATCCGCGAGCTGTGATCGAGTGAAGCATATCGACCACTTCGCGGAGAGAGGTC 132  
 Qy 43 LysPheGlyAlaGlyIleGlyTrpIleArgGluGluMetGlyIleSerAsnSerProAspIle 62

Db 133 AAATTCGCGCGCGCAAGGTATCGCGACGCGCATGGCGGCGAGGCGAGATGCTGGCGCC 192  
 Qy 63 AsnThrIleAspLeuValIleThrAsnAlaMetIleIleAspTrpThrGlyIleTrpIle 82  
 Db 193 GACTGTGTGACCTGT 252  
 Qy 83 AlaAspIleGlyIleIleValAsnGlyIleIleIleGlyIleIleGlyIleIleGlyIleIle 102  
 Db 253 GCCGATTCGCGCGTGAAGGAGACGCGCGGATCTTCGCGCATTCGCGCAAGCGGTAACCCGAT 312  
 Qy 103 MetGlnAspGlyValSerProHisMetValAlaGlyValGlyIleThrGluAlaLeuAlaGly 122  
 Db 313 ATCCAGGCCCAACGTCAC-----ATCCCATCGGTGCTGTGACGGAAGTATGCGCGCC 366  
 Qy 123 GluGluMetIleIleThrAlaGlyIleIleAspSerHisThrHisPheLeuSerProGln 142  
 Db 367 GAAAGAAAATGTGTACCGCGCGGAGATCGATCCCATATTCATCTGGATGTGTCCGAG 426  
 Qy 143 GlnPheProThrAlaLeuAlaAsnGlyValIleThrIleMetPheGlyIleGlyIleThrGlyPro 162  
 Db 427 CAGCGGAAAGGCGCTGT 486  
 Qy 163 ValAspGlyThrAsnAlaThrThrIleThrProGlyIleTrpAsnLeuHisArgMetLeu 182  
 Db 487 GCGCGGCGCACCATGCGCACACCATGCGACCCCGCGCGGTGTATCTCGCATGCGATGCTG 546  
 Qy 183 ArgAlaAlaGluGluTrpSerMetAsnValGlyIleGlyIleGlyIleGlyIleGlyIle 202  
 Db 547 CAGCGCGCGCACCGCTGCGCATATTCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 606  
 Qy 203 LysIleGlnIleValGluGlnValGluAlaGlyValIleGlyIlePheIleValIleGlyAsp 222  
 Db 607 CCGAGTGCCTGCGCGGACAGGTGCGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 666  
 Qy 223 TrpGlyThrThrProSerAlaIleAspHisCysLeuSerValAlaAspGlyTrpAspVal 242  
 Db 667 TGGGGCGGACCCCGGGGGGATGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 726  
 Qy 243 GlnValCysIleHisThrAspThrValAsnGluAlaGlyTrpValAspAspThrLeuAsn 262  
 Db 727 CAGGTGCGCGTGCACACGACGACACCTTAATGAATCCGTTTGTGTGTGTGTGTGTGTGTGT 786  
 Qy 263 AlaMetAsnGlyAspAlaIleHisAlaIleIleIleIleIleIleIleIleIleIleIleIle 282  
 Db 787 GCGATCGCGCGCGCACATTCACACCTTCATACCAAGGCGCGCGCGCGCGCGCGCGCGCG 846  
 Qy 283 ProAspValIleThrMetAlaGlyIleLeuAsnIleLeuProSerSerThrThrProThr 302  
 Db 847 CCGGACATCATCACCGCTGCG 906  
 Qy 303 IleProTrpThrIleAsnThrValAlaGluHisLeuAspMetLeuMetThrCysHisHis 322  
 Db 907 CTGCGCTTACACCTTCACACCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 966  
 Qy 323 LeuAspIleArgIleIleGlyIleAspLeuGlnPheSerIleIleArgIleIleArgProGlySer 342  
 Db 967 CTGACCCGCGACATCGCGGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1026  
 Qy 343 IleAlaIleAlaGluAspValIleHisAspMetGlyValIleAlaMetThrSerSerAspSer 362  
 Db 1027 ATGCGTGGGAAAGCGT 1086  
 Qy 363 GlnAlaMetGlyArgAlaGlyIleValIleProArgThrTrpGlnThrAlaAspIleAsn 382  
 Db 1087 CAGGCGATGGGCGCGT 1146  
 Qy 383 LysIleGluPheGlyIleIlePheProGluAspGlyIleAspAsnAspAsnAspAsnAsp 402  
 Db 1147 AAGGTGACGCGCGAGCGT 1206  
 Qy 403 ArgTrpIleSerIleTrpThrIleAsnProAlaLeuThrHisGlyValSerGlyTrpIle 422

Db	1207	CGCTACATCGCGAAATACACCATCAACCCGGCGCTGACCCAGGCATGGACACGAAGTC	1266
Qy	423	GLYSERVALGIUGIUGIYLYSEIIEALASPHEUVALVATYPASPProIALPhePheGly	442
Db	1267	GGATTCATTGGAGGTGGGTAACTGGCTGAACCTCGGTGTGTGTACACGACCTTCTTGGCG	1326
Qy	443	VALYSPProLYSILEVALIIEIYSGIYGLYKETVALVALPheSERGIUMETGLYASPSER	462
Db	1327	GTEAAACCGCGCACCGTGTATAAAGCGCGCATGATCGGCATCGCGCGCATGGCGCATTC	1386
Qy	463	ASNAIAsERVALProTHrProGInProVALTYTYTARGIUMETPheGlyHISGly	482
Db	1387	AATGCACTTATTCGCAACCCCGACGCGGTGACATACCGCCGAGTGTGGCGCGCTGGGC	1446
Qy	483	LYSALYSPhEASpTHSerILEThPheVALSerYSVALAIATYGIUASNGLYVAL	502
Db	1447	AGCGCCCGCATCACTCGCGCTCACTCTCTGTGCGAGGCGCGCGAGCCAAATGGCCTT	1506
Qy	503	LYSGIUUSLEUGIYLEUGIUAIRGINVALIeUProVALIYSAANCYSARGASNIIEThr	522
Db	1507	GCCGAGCGGCTGAACCTGCGGACGCGCATCGCCGNGTGAAGAAAGGTGCGTACAGTGCGAG	1566
Qy	523	LYSLYSAspPHELYSPheASNApLYSThrALALYSILEThrVALASPProLYSThrPhe	542
Db	1567	AAAGCCGACATGTGCACACACAGCGCTGACGCTTAACATCACCTGTGAGGCCAAGACTAT	1626
Qy	543	GLIUVAlPheVALASPGLYLYSLEUCYSThrSERIYSProTHrSERGINVALProLEUAla	562
Db	1627	GAGGTGCGGGTGGATGGCAACTTATATCACAGCGAGCGGAGACAGTCCTTCGCGATGGCG	1686
Qy	563	GLINARGTYR 565	
Db	1687	CAACGATAT 1695	

Db 124 GCGCAAGAGGCTCAATTGGTGTGGTGTAAAGTATTTCGTGATGCGTATGGGGCAACCCAA 183  
 QY 59 SerProAspGluAsnThrLeuAspLeuValIleThrAsnAlaMetIleLeaPylrThr 78  
 Db 184 GTTGTAACTGCTGAGTGTGTCATGTCATTCGATCCACCAAGCCATATTTTAAATTATTGG 243  
 QY 79 GlyIleTyrlValAlaAspIleGlyIleLeuAsnGlyLysIleHisIleGlyIleGlyLysAla 98  
 Db 244 GGCATTGTAAAGCAGATTTGGCATTAAGAAGTGGCCGATTTGTCCGTATTTGCCAAGCA 303  
 QY 99 GlyAsnLysAspMetGlnAspGlyValSerProHisMetValAlaGlyValGlyThrGlu 118  
 Db 304 GGATATCCAGATGTTCAGGCCCAATGTGAT-----ATTGTCATTGGCCCGGAAACAGAA 357  
 QY 119 AlaLeuAlaGlyGluGlyMetIleLeuThrAlaGlyGlyIleAspSerHisThrHisPhe 138  
 Db 358 GTTGTGGCGGGGAAGGTAAATGTGCTGCTGGTGGTGGATTTGATACCATATCCACTTT 417  
 QY 139 LeuSerProGlnGlnPheProThrAlaLeuAlaAsnGlyValaThrThrMetPheGlyIly 158  
 Db 418 ATTGTCGCGCAACAGACCCAAAGAGGTGCTGTTTCTGGGGTAAACCACTTTATTTGGTGA 477  
 QY 159 GlyThrGlyProValAspGlyThrAsnAlaThrThrIleSerProGlyLysThrAsnLeu 178  
 Db 478 GGAAACAGGCCCTGTGGCGGTACTATGCAACCAACGTTAACCCCGGTATTGGAAATATG 537  
 QY 179 HisArgMetLeuArgAlaIleGluGluIuTyrlSerMetAsnValGlyPheLeuGlyLysGly 198  
 Db 538 TACCGCATGTTAAGGCGGTGATGAATTACCATTAATGTGGGGTTATTTGGTAAAGT 597  
 QY 199 LeuSerSerSerLysAlaGluLeuValGluGluGluValAlaGlyValAlaIleGlyPheLys 218

RESULT 15  
 US-09-543-681A-1857  
 Sequence 1857, Application US/09543681A  
 Patent No. 6605709  
 GENERAL INFORMATION:  
 APPLICANT: GARY BRETTON  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
 TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
 FILE REFERENCE: 2709.1002-001  
 CURRENT APPLICATION NUMBER: US/09/543, 681A  
 CURRENT FILING DATE: 2000-04-05  
 PRIOR APPLICATION NUMBER: US 60/128,706  
 PRIOR FILING DATE: 1999-04-09  
 NUMBER OF SEQ ID NOS: 8344  
 SEQ ID NO 1857  
 LENGTH: 1710  
 TYPE: DNA  
 ORGANISM: Proteus mirabilis  
 US-09-543-681A-1857

[illegible]

Alignment Scores:		
Pred. NO.:	1,366-193	Length: 1710
Score:	1818.50	Matches: 340
Percent Similarity:	75.44%	Conservative: 90
Best Local Similarity:	59.65%	Mismatches: 135
Query Match:	60.64%	Indels: 5
DB:	4	Gaps: 2
US-09-904-994B-3 (1-568) x US-09-543-681A-1857 (1-1710)		

QY 319 ThrCysHisHisLeuAspLysArgIleArgGluAspLeuGlnPheSerGlnInsertArgIle 338  
Db 958 GCTCTGATCATCTCGATCCCTCTATTCTCTGAAGATGCGCATTTCTGATATCTCGATT 1010  
QY 339 ArgProGlySerIleAlaIaGluAspValLeuHisAspMetGlyValIleAlaMetThr 358  
Db 1018 CCGTCGAAACCATTCCTGCAAGAAGTATCTTACATGATATGGGGCAATTTGGTGATG 1070

```

QY 2 LysMetLys-----LysGInGluTrpValAsnThrCysGlyProThrLysGlyAsp 18
Db 4 AAAATGAAACTACTCTCACTGACAGCTTATGCGGATTTGTTGGCCCAACACAGCGCAT 63
   19 LysValAlaArgLeuGlyAspThrAspLeuTrpAlaGluValGluIleAspTrpThrTrp 38
   64 CGTTTGGCATTAAGACGATACCGAGCTTTCTTGAAATTAATAAAGATTTCCACACTTAT 12
QY 39 GlyGluGluLeuLysPheGlyAlaGlyLysThrIleArgGluGluLysMetGlyGlnSerAsn 58

```

Db 1078 TCGTCAGACTCACAGCCATTGGGACGACGTCCGAGAAATTATCTTTAGCAGCTTGCGACGTCT 113

QY 379 A1aaaplysaasnllyesglubphegtlylsleuProgluaapcllylsapapaanaapAsn 399

Db 1138 GCACATAAATATGAATTCGAACGAGGACATTGACGGGTGATATGCCGACGATTAATGATTAAT 115

QY 399 pheargilelyeargtyrllleserlytyrrhllsaasproalaubauthrhlsqlyval 418

```
Db 1198 AATCGATTAAACGTTATATGCTAAATACAGATTAAATCCCGCACTGGCAGATGGCAATT 1257
QY SerGluTyrIleGlySerValGluGluGlyLeuIleAlaAspLeuValTyrAspPro 438
Db 1258 GCTCATACGGGTGGATCATAGAAAAGGTAACCTTGGATATCGTCTATGGATCCT 1317
QY AlaPhePheGlyValLysProLysIleValIleLysGlyGlyMetValLysPheSerGlu 458
Db 1318 GCTTCTTTGGCGTCAAAACCGCACTTATCATAAAGGTGGATGGTCTTATCGGCCA 1377
QY MetGlyAspSerAsnAlaSerValProThrProGlnProValTyrTyrArgGluMetPhe 478
Db 1378 ATGGGGGATATTAAATGCGGCTATTCCAACCGCAACCGGTTCAATTATCGTCCAATGTAT 1437
QY GlyHisHisGlyLysAlaLysPheAspPheSerIleThrPheValSerLysValAlaTyr 498
Db 1438 GCCTGTTTAGGAAAAGCCAAATATCAACGTGATGATCTTATGTCAAAAGCGGTATT 1497
QY GluAsnGlyValLysGluLysLeuGlyLeuGluArgGlnValLeuProValLysAsnCys 518
Db 1498 GAGCGGGAGTGCAGAAAATTTAGGCTTAATAAGCTTAATTGTCGTGTGGAGGCTGT 1557
QY ArgAsnIleThrLysLysAspPheLysAsnAspLysThrAlaLysIleThrValAsp 538
Db 1558 CGTCATATCACAAAAGCTTCGATGATCCACATATACTATGTTCTCATATCGAATTAGAT 1617
QY ProLysThrPheGluValPheValAspGlyLysLeuCysThrSerLysProThrSerGln 558
Db 1618 CCACAAACTTACATGTTAAAGCGATGGTGTACCACTGTTGTGAGCCAGCACTGAA 1677
QY ValProLeuAlaGlnArgTyrThrPhePhe 568
Db 1678 TTACCGATGGCTCAACGCTATTCTTATT 1707
```

Search completed: February 16, 2005, 01:31:21  
Job time : 2119.33 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 15, 2005, 20:34:30 : Search time 8.53904 Seconds  
(Without alignments)  
1975.712 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKEQKFLLYAGEVA.....ADSRHKHKAJDKXKSHQFIK 226

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 324380

Minimum DB seq length: 0  
Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281.5	23.8	100	1	US-08-967-513-3
2	281.5	23.8	100	2	US-08-687-645B-3
3	269.5	22.8	99	1	US-07-732-342C-1
4	261.5	22.1	100	2	US-08-467-822-23
5	261.5	22.1	100	3	US-08-432-697-23
6	261.5	22.1	100	3	US-08-466-248-23
7	250.5	21.2	100	3	US-09-602-777A-18
8	68.5	5.8	40	4	US-09-602-777A-12
9	62	5.2	97	4	US-09-640-211A-973
10	60	5.1	63	4	US-09-248-796A-27427
11	58	4.9	18	4	US-08-945-038-12
12	57	4.8	83	4	US-09-270-767-60797
13	57	4.8	92	4	US-09-513-999C-4321
14	57	4.8	100	4	US-09-328-352-7014
15	56.5	4.8	83	4	US-09-248-796A-23863
16	56	4.7	15	3	US-09-081-001-1
17	56	4.7	64	4	US-09-328-352-6741
18	56	4.7	86	4	US-09-621-976-4833
19	55.5	4.7	100	4	US-09-621-976-7112
20	54.5	4.6	94	4	US-09-270-767-36087
21	54.5	4.6	94	4	US-09-270-767-51304
22	54	4.6	72	4	US-09-540-236-3389
23	54	4.6	85	4	US-09-248-796A-28048
24	53	4.5	82	4	US-09-689-065B-8
25	52	4.4	77	3	US-08-936-165A-511
26	51.5	4.4	65	4	US-09-248-796A-26949
27	51.5	4.4	68	4	US-09-107-433-4410

28	51.5	4.4	72	4	US-09-512-563C-61	Sequence 61, Appl
29	51.5	4.4	81	1	US-08-225-757B-8	Sequence 8, Appl
30	51	4.3	66	4	US-09-732-210-881	Sequence 881, App
31	51	4.3	100	4	US-09-540-236-3479	Sequence 3479, App
32	50.5	4.3	80	4	US-09-252-991A-25293	Sequence 25293, A
33	50.5	4.3	90	4	US-09-902-540-10059	Sequence 10059, A
34	50.5	4.3	93	4	US-09-248-796A-26429	Sequence 26429, A
35	50.5	4.3	94	3	US-09-147-550-81	Sequence 81, Appl
36	50.5	4.3	94	3	US-09-557-917-81	Sequence 81, Appl
37	50	4.2	86	4	US-09-621-976-6687	Sequence 6687, App
38	49.5	4.2	45	4	US-09-079-030-112	Sequence 112, App
39	49.5	4.2	85	4	US-09-134-000C-3459	Sequence 3459, App
40	49.5	4.2	86	4	US-09-543-681A-5320	Sequence 5320, App
41	49.5	4.2	94	3	US-09-147-550-38	Sequence 38, Appl
42	49.5	4.2	94	3	US-09-557-917-38	Sequence 38, Appl
43	49.5	4.2	95	4	US-09-513-999C-5808	Sequence 5808, App
44	49	4.1	79	4	US-09-513-999C-6886	Sequence 6886, App
45	49	4.1	83	3	US-09-138-721-19	Sequence 19, Appl

#### ALIGNMENTS

RESULT 1  
US-08-967-513-3  
Sequence 3, Application US/08967513  
Patent No. 5783436  
GENERAL INFORMATION:  
APPLICANT: Robert P. Hausinger  
TITLE OF INVENTION: Mutant Urease and Method  
TITLE OF INVENTION: of Use For Determination  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/967,513  
FILING DATE: 11-NOV-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/687,645  
FILING DATE: July 26, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: MSU 4.1-309  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 5783436e  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100  
TYPE: amino acids  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: NO  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: N/A

STRAIN: N/A  
INDIVIDUAL ISOLATE: N/A  
CELL TYPE: N/A  
FEATURE:  
NAME/KEY: subunit Urea  
LOCATION:  
IDENTIFICATION METHOD: Sequencing  
OTHER INFORMATION: encoded subunit of  
OTHER INFORMATION: mutant urease  
US-08-967-513-3

Query Match 23.8%; Score 281.5; DB 1; Length 100;  
Best Local Similarity 59.0%; Pred. No. 4e-24;  
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGVARRKKAEGKLNQPEALAIYISAHIMDEARCKTTVAQLMEE 60  
DB 1 MELTPREKDKLLFTALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59  
QY 61 CMHPLKDEVMPCVGMV/PDLGVEATPPDGTCLVTNNMPI 100  
DB 60 GRHVLTRQVMEGVPEMIPDIOVEATPPDGSCLVTNNMPI 99

RESULT 2  
US-08-687-645B-3  
Sequence 3, Application US/08687645B  
Patent No. 5846752

GENERAL INFORMATION:  
APPLICANT: Robert P. Hausinger  
TITLE OF INVENTION: Mutant Urease and Method  
TITLE OF INVENTION: of Use For Determination  
TITLE OF INVENTION: of Urea  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,645B  
FILING DATE: July 26, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:

FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: MSU 4.1-309  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
TELEX: No. 5846752e  
INFORMATION FOR SEQ. ID NO. 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100  
TYPE: amino acids  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: No  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:

ORGANISM: N/A  
STRAIN: N/A  
INDIVIDUAL ISOLATE: N/A  
CELL TYPE: N/A  
FEATURE:  
NAME/KEY: subunit Urea  
LOCATION:  
IDENTIFICATION METHOD: Sequencing  
OTHER INFORMATION: encoded subunit of  
OTHER INFORMATION: mutant urease  
US-08-687-645B-3

Query Match 23.8%; Score 281.5; DB 2; Length 100;  
Best Local Similarity 59.0%; Pred. No. 4e-24;  
Matches 59; Conservative 18; Mismatches 22; Indels 1; Gaps 1;

QY 1 VKLTPEQEKFLYYAGVARRKKAEGKLNQPEALAIYISAHIMDEARCKTTVAQLMEE 60  
DB 1 MELTPREKDKLLFTALVAERRLARGLKLNYPESVALISAFIMEGARDG-KSVASLMEE 59  
QY 61 CMHPLKDEVMPCVGMV/PDLGVEATPPDGTCLVTNNMPI 100  
DB 60 GRHVLTRQVMEGVPEMIPDIOVEATPPDGSCLVTNNMPI 99

RESULT 3  
US-07-732-242C-1  
Sequence 1, Application US/07732242C  
Patent No. 5298399

GENERAL INFORMATION:  
APPLICANT: Uozumi, Takeshi; Maeaki, Haruhiko;  
APPLICANT: Hidaka, Makoto; Nakamura, Akira;  
APPLICANT: Maeda, Michihisa; Yoneta, Yasuo  
TITLE OF INVENTION: Gene of Urease  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Frlshauf, Holcz, Goodman & Woodward, P. C.  
STREET: 600 Third Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10016-2088

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 0.72mb  
COMPUTER: IBM PC compatible (NEC PC-9801 RX)  
OPERATING SYSTEM: MS DOS  
SOFTWARE: ASCII Form  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/732,242C  
FILING DATE: 19910718  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JPN 2-210178  
FILING DATE: 10-AUG-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Goodman, Herbert  
REGISTRATION NUMBER: 17081  
REFERENCE/DOCKET NUMBER: 910532/HG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 972-1400  
TELEFAX: (212) 370-1622  
TELEX: 236268

INFORMATION FOR SEQ. ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 99 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-732-242C-1

Query Match 22.8%; Score 269.5; DB 1; Length 99;  
Best Local Similarity 55.6%; Pred. No. 9e-23;

Matches 55; Conservative 18; Mismatches 25; Indels 1; Gaps 1;  
Qy 2 KLTPEOEKFLYYAGVARRRKAEGIKLNOPEAIYISAHIMDEARRGKTTVAQLMEEC 61  
Db 1 KLTREMEKLMIVAAADLARRRERGLKLNYPNAVAMITYEVLGADG-KTVAQLMQYG 59  
Qy 62 MHLFKDEVMPCGVGNV/PDLGVEATPPDGTCLVTVNMP1 100  
Db 60 ATILTKEQVMEGVAMIPDIOIEATFPDGTCLVTVHDP1 98

RESULT 4  
US-08-467-822-23  
; Sequence 23, Application US/08467822  
; Patent No. 5843460  
; GENERAL INFORMATION:  
; APPLICANT: Labigne, Agnes  
; APPLICANT: Sauerbaum, Sebastien  
; APPLICANT: Ferrero, Richard L.  
; APPLICANT: Thiberge, Jean-Michel  
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
; TITLE OF INVENTION: POLYPEPTIDES  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fiumegan, Henderson, Farabow, Garrett &  
; ADDRESSER: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,822  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/447,177  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/432,697  
; FILING DATE: 02-MAY-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0137-02000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-467-822-23

Query Match 22.1%; Score 261.5; DB 2; Length 100;  
Best Local Similarity 56.0%; Pred. No. 7.4e-22;  
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;  
Qy 1 VKLTPEOEKFLYYAGVARRRKAEGIKLNOPEAIYISAHIMDEARRGKTTVAQLMEE 60  
Db 1 MELTPREKDLILFTALVARRRRLAKGLKLNYPNAVAMITYEVLGADG-KTVAQLMSE 59

Qy 61 CMHFLKDEVMPCGVGNV/PDLGVEATPPDGTCLVTVNMP1 100  
Db 60 GRTVLTAEQVMEGVAMIPDIOIEATFPDGTCLVTVHDP1 99

RESULT 5  
US-08-432-697-23  
; Sequence 23, Application US/08432697  
; Patent No. 6248330  
; GENERAL INFORMATION:  
; APPLICANT: Labigne, Agnes  
; APPLICANT: Sauerbaum, Sebastien  
; APPLICANT: Ferrero, Richard L.  
; APPLICANT: Thiberge, Jean-Michel  
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
; TITLE OF INVENTION: POLYPEPTIDES  
; NUMBER OF SEQUENCES: 44  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fiumegan, Henderson, Farabow, Garrett &  
; ADDRESSER: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/432,697  
; FILING DATE: 02-MAY-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyers, Kenneth J.  
; REGISTRATION NUMBER: 25,146  
; REFERENCE/DOCKET NUMBER: 03495.0137-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 408-4400  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 100 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-432-697-23

Query Match 22.1%; Score 261.5; DB 3; Length 100;  
Best Local Similarity 56.0%; Pred. No. 7.4e-22;  
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;  
Qy 1 VKLTPEOEKFLYYAGVARRRKAEGIKLNOPEAIYISAHIMDEARRGKTTVAQLMEE 60  
Db 1 MELTPREKDLILFTALVARRRRLAKGLKLNYPNAVAMITYEVLGADG-KTVAQLMSE 59  
Qy 61 CMHFLKDEVMPCGVGNV/PDLGVEATPPDGTCLVTVNMP1 100  
Db 60 GRTVLTAEQVMEGVAMIPDIOIEATFPDGTCLVTVHDP1 99

RESULT 6  
US-08-466-248-23  
; Sequence 23, Application US/08466248  
; Patent No. 6258359  
; GENERAL INFORMATION:  
; APPLICANT: Labigne, Agnes  
; APPLICANT: Sauerbaum, Sebastien

```

; APPLICANT: Ferrero, Richard L.
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,248
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 100 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-466-248-23

Query Match      22.1%; Score 261.5; DB 3; Length 100;
Best Local Similarity 56.0%; Pred. No. 7.4e-22;
Matches 56; Conservative 17; Mismatches 26; Indels 1; Gaps 1;

QY      1 VKTTPGEKFLIYAGVARKKABGLKNOPEALAYISAHIMDEARRGKTVTAQLMEER 60
      :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      1 MELTPREKDLTLFTAGLVALERLAKGLKLNPERVALISCAIMEGARREG-KTVAQLMSB 59

QY      61 CMHFLKDEVMVPGVNMVPLDGEATFPDGTKLVTVMWPI 100
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      60 GRTVLTAEQVMEGVPEMINKDQVECTFPDGTKLVTSHSPI 99

RESULT 7
US-09-602-777A-18
; Sequence 18, Application US/09602777A
; Patent No. 6831165
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroege, Burkhard
; APPLICANT: Schröder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habethauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN HOMEOSTASIS AND ADAPTATION
```

```

; FILE REFERENCE: BGI-128CP
; CURRENT APPLICATION NUMBER: US/09/602,777A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: DE 19931636.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932125.6
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932126.4
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932127.2
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932128.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932129.9
; PRIOR FILING DATE: 1999-07-19
; PRIOR APPLICATION NUMBER: DE 19932226.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932920.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932922.2
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932924.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932928.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932930.3
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932933.8
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932935.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19932973.7
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933002.6
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933003.4
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933005.0
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19933006.9
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19941378.9
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941390.8
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19941391.6
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: DE 19942088.2
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 442
; SEQ ID NO 18
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-602-777A-18

Query Match      21.2%; Score 250.5; DB 4; Length 100;
Best Local Similarity 52.0%; Pred. No. 1.3e-20;
Matches 51; Conservative 20; Mismatches 26; Indels 1; Gaps 1;

QY      3 LTPGEKFLIYAGVARKKABGLKNOPEALAYISAHIMDEARRGKTVTAQLMEECM 62
      :::::::::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      3 ITPREKLMIVVADLARRRKDRGLKLNHPAVALITYELLEGARDG-RTVADLMSGS 61

QY      63 HFLKDEVMVPGVNMVPLDGEATFPDGTKLVTVMWPI 100
      ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB      62 TLTTRDVLAEGLPEMIPDQVEATFPDGTKLVTVMWPI 99
```



RESULT 8  
US-09-602-777A-12  
Sequence 12, Application US/09602777A  
Patent No. 6831165  
GENERAL INFORMATION:  
APPLICANT: Pompejun, Markus  
APPLICANT: Krieger, Burkhard  
APPLICANT: Schroder, Hartwig  
APPLICANT: Zeider, Oskar  
APPLICANT: Haberhauser, Gregor  
TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
FILE REFERENCE: BGI-128CP INVOLVED IN HOMEOSTASIS AND ADAPTATION  
CURRENT APPLICATION NUMBER: US/09/602,777A  
CURRENT FILING DATE: 2000-06-23  
PRIOR APPLICATION NUMBER: US 60/141031  
PRIOR FILING DATE: 1999-06-25  
PRIOR APPLICATION NUMBER: DE 19931636.8  
PRIOR FILING DATE: 1999-07-08  
PRIOR APPLICATION NUMBER: DE 19932125.6  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932126.4  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932127.2  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932128.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932129.9  
PRIOR FILING DATE: 1999-07-19  
PRIOR APPLICATION NUMBER: DE 19932226.0  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: DE 19932920.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932922.2  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932924.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932928.1  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932930.3  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932933.8  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932935.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19932973.7  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933002.6  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933003.4  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933005.0  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19933006.9  
PRIOR FILING DATE: 1999-07-14  
PRIOR APPLICATION NUMBER: DE 19941378.9  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941379.7  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941390.8  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19941391.6  
PRIOR FILING DATE: 1999-08-31  
PRIOR APPLICATION NUMBER: DE 19942088.2  
PRIOR FILING DATE: 1999-09-03  
NUMBER OF SEQ ID NOS: 442  
SEQ ID NO 12  
LENGTH: 40  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-602-777A-12

Query Match 5.8%, Score 68.5; DB 4; Length 40;

Best Local Similarity 31.1%; Pred. No. 1.5;  
Matches 14; Conservative 8; Mismatches 16; Indels 7; Gaps 1;  
Cy 102 PDEHFKAGEVFGCDKDIENLAKGEVTELEVTNEGPKSLHVGSHF 146  
Db 3 FGEYILSESLTG-----NVGREAKTIEIINTGDRPVQGSHF 40

RESULT 9  
US-09-640-211A-973  
Sequence 973, Application US/09640211A  
Patent No. 6833446  
GENERAL INFORMATION:  
APPLICANT: Wood, Marion  
APPLICANT: Shenk, Michael A.  
APPLICANT: McGrath, Annette  
APPLICANT: Glenn, Matthew  
TITLE OF INVENTION: Compositions and Methods for the  
TITLE OF INVENTION: Modification of Gene Transcription  
FILE REFERENCE: 11000.1021CIU  
CURRENT APPLICATION NUMBER: US/09/640,211A  
CURRENT FILING DATE: 2000-08-16  
NUMBER OF SEQ ID NOS: 2368  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 973  
LENGTH: 97  
TYPE: PRT  
ORGANISM: Pinus radiata  
US-09-640-211A-973

Query Match 5.2%; Score 62; DB 4; Length 97;  
Best Local Similarity 31.0%; Pred. No. 32;  
Matches 18; Conservative 14; Mismatches 18; Indels 8; Gaps 3;

Cy 114 GCDKDIIE-LNAGKEVTELEVTNEGPKSLHVGSHFHFPEANKALKPDRKAYGRLDI 169  
Db 39 GCERKVNALNMSMGVSVVERQYKATVTG-----YVDANKVLK--RVKQTGKKAL 90

RESULT 10  
US-09-248-796A-27427  
Sequence 27427, Application US/09248796A  
Patent No. 6747137  
GENERAL INFORMATION:  
APPLICANT: Keith Weinstock et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS  
FILE REFERENCE: 107196.132  
CURRENT APPLICATION NUMBER: US/09/248,796A  
CURRENT FILING DATE: 1999-02-12  
PRIOR APPLICATION NUMBER: US 60/074,725  
PRIOR FILING DATE: 1998-02-13  
PRIOR APPLICATION NUMBER: US 60/096,409  
PRIOR FILING DATE: 1998-08-13  
NUMBER OF SEQ ID NOS: 28208  
SEQ ID NO 27427  
LENGTH: 63  
TYPE: PRT  
ORGANISM: Candida albicans  
US-09-248-796A-27427

Query Match 5.1%; Score 60; DB 4; Length 63;  
Best Local Similarity 30.9%; Pred. No. 28;  
Matches 17; Conservative 9; Mismatches 25; Indels 4; Gaps 2;  
Cy 168 DIPSGNTLRIGAGQTRKVLIPLGSKKVTGMNGLVNNIDERRHKALPKAKSH 222  
Db 11 DIASVVGISGKNV--ADKLHLGGVGHIVG--GVAGGIAGNLEHKADEBEVREH 61

RESULT 11  
US-08-945-038-12  
Sequence 12, Application US/08945038

```

Patent No. 6762295
GENERAL INFORMATION:
APPLICANT: Doidge, Christopher V.
APPLICANT: Lee, Adrian
APPLICANT: Radcliff, Fiona J.
APPLICANT: Hocking, Diana M.
APPLICANT: Webb, Elizabeth A.
TITLE OF INVENTION: PROTECTIVE HELICOBACTER ANTIGENS
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/945,038
FILING DATE: 23-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU96/00225
FILING DATE: 19-APR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN2575
FILING DATE: 21-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN3931
FILING DATE: 03-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PN7565
FILING DATE: 16-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 017227/0133
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
US-08-945-038-12

Query Match 4.9%; Score 58; DB 4; Length 18;
Best Local Similarity 61.1%; Pred. No. 7;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 VKLTPKEOEKFLVYAGE 18
:|||||:|:|:|
Db 1 MKLTPKELDKMLHRAGE 18

RESULT 12
US-09-270-767-60797
Sequence 60797, Application US/09270767
Patent No. 6703491
GENERAL INFORMATION:
APPLICANT: Homburger et al.
TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
FILE REFERENCE: File Reference: 7326-094
CURRENT APPLICATION NUMBER: US/09/270,767
CURRENT FILING DATE: 1999-03-17

```

```

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 60797
; LENGTH: 83
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-60797

Query Match          4.8%; Score 57; DB 4; Length 83;
Best Local Similarity 28.6%; Pred. No. 92;
Matches 24; Conservative 13; Mismatches 37; Indels 10; Gaps 3;

OY      40 SAHIMDEARRGKKTVAQLMEBCMHFLPKDEYVPGVGMVDP----LGVEATFPDGTKLVT 95
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1 AAQAAASADGATITISELMTKTVAFHAPGENFKAQDGYVTEHTERLLKEHLARTGKQVHT 60

OY      96 VMNPIDPD----EHFKAQGEVKEFG 114
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      61 -RPPPEPGLIHGAKAININFG 83

RESULT 13
US-09-513-999C-4321
; Sequence 4321, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.Pm
; SEQ ID NO 4321
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -20...-1
; OTHER INFORMATION: score 3.5
; OTHER INFORMATION: seq IVLGTGLTCLTS/GI
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 43
; OTHER INFORMATION: Xaa=Pro or Thr
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 51
; OTHER INFORMATION: Xaa=Asp or Glu
US-09-513-999C-4321

Query Match          4.8%; Score 57; DB 4; Length 92;
Best Local Similarity 30.6%; Pred. No. 1,1e+02;
Matches 19; Conservative 3; Mismatches 22; Indels 18; Gaps 2;

OY      153 KALKFDEKAYG-----KSLDIPSGNTLRIGAGQTRKVLIPYGGSKKVIYM 199
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      29 KVLHMDRNPYYGGSASITPLEDLYKRFKIFGSPXESMGGRKXNVDLIP-----KFLMA 83

OY      200 NG 201
      : :
Db      84 NG 85

RESULT 14
US-09-328-352-7014
; Sequence 7014, Application US/09328352

```

```

: Patent No. 6562958
: GENERAL INFORMATION:
: APPLICANT: Gary L. Breton et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
: TITLE OF INVENTION: BAUMANNIT FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC99-03PA
: CURRENT APPLICATION NUMBER: US/09/328,352
: CURRENT FILING DATE: 1999-06-04
: NUMBER OF SEQ ID NOS: 8252
: SEQ ID NO 7014
: LENGTH: 100
: TYPE: PR1
: ORGANISM: Acinetobacter baumannii
: US-09-328-352-7014

```

Query Match	4.8%	Score	57	DB	4	Length	100
Best Local Similarity	25.0%	Pred.	NO	1.2e+02			
Matches	11	Conservative	10	Mismatches	23	Indels	0
						Gaps	0

QY 111 VKFGCDKDIELNAGKEVTELEVTNEGPKSLHVGSHPHFFPEANKA 154  
 Db 20 VELTCESYIHIPSGASGNKLGIDNRAEDLIHITEDPSFPKTEA 63

```

RESULT 15
US-09-248-796A-23863
: Sequence 23863 Application US/09248796A
: Patent No. 6747137
: GENERAL INFORMATION:
: APPLICANT: Keith Weinstock et al
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
: TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.132
: CURRENT APPLICATION NUMBER: US/09/248,796A
: CURRENT FILING DATE: 1999-02-12
: PRIOR APPLICATION NUMBER: US 60/074,725
: PRIOR FILING DATE: 1998-02-13
: PRIOR APPLICATION NUMBER: US 60/096,409
: PRIOR FILING DATE: 1998-08-13
: NUMBER OF SEQ ID NOS: 28208
: SEQ ID NO 23863
: LENGTH: 83
: TYPE: PRT
: ORGANISM: Candida albicans
: US-09-248-796A-23863

```

Query Match	4.8%	Score	56.5	DB	4	Length	83
Best Local Similarity	31.8%	Pred. No.	1.1e+02				
Matches	14	Conservative	8	Mismatches	19	Indels	3
						Gaps	1

QY 122 NAGKEVTELEVNTNEGPKSLHVG---HFHFFEANKALKKFDREKA 162  
| | | : : : : : | | | | |  
Db 4 NQGEIERIKIVAKRPTCVTIGSGIKNNEMFVTTHLKLDEKA 47

Search completed: February 15, 2005, 20:44:31  
Job time : 9.53904 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus model

Run on: February 15, 2005, 20:44:36 ; Search time 87.6675 Seconds

(without alignments)  
4218.194 Million cell updates/sec

Title: US-09-904-994B-2

Perfect score: 1181

Sequence: 1 VKLTPKQKFKLYVAGEVA.....ADERRHKALDKAKSHGRFK 226

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Xgapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81818359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame+g2n.model -DEV=x1h  
-Q=/cgnt\_1/USPTO.epool/US09904994/runat\_11022005\_145648\_25560/app\_query.fasta\_1.1102  
-DB=Issued\_Patents\_NA -QMT=faeap -SUFFIX=rn1 -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdd  
-LIST=45 -DOCALLIGN=200 -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MABLEN=2000000000  
-USER=US09904994 -CGCN 1.1.77 -@runat\_11022005\_145648\_25560 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued\_Patents\_NA.\*  
1: /cgnt\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgnt\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgnt\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgnt\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgnt\_6/ptodata/1/ina/PTCUS.COMB.seq.\*  
6: /cgnt\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	610.5	51.7	2735	2	US-08-920-095-1
2	610.5	51.7	2735	2	PCT-US96-05800-1
3	610.5	51.7	4824	4	US-09-431-705-1
4	610.5	51.7	4824	4	US-09-431-705-19
5	572.5	48.5	2619	2	US-08-467-822-19
6	572.5	48.5	2619	2	US-08-433-697-19
7	572.5	48.5	2619	3	US-08-466-248-19
8	484.5	41.9	3475	4	US-09-710-279-4032
9	480.5	40.7	5966	4	US-08-956-171E-22
10	480.5	40.7	5966	4	US-08-781-966A-22
11	474	40.1	2400	1	US-08-967-513-1
12	474	40.1	2400	2	US-08-687-645B-1

13	468.5	39.7	6131	1	US-07-732-242C-8	Sequence 8, Appli
14	467	39.5	1830121	4	US-09-557-884-1	Sequence 1, Appli
15	467	39.5	1830121	4	US-09-643-890A-1	Sequence 1, Appli
16	445.5	37.7	8729	3	US-09-453-702B-258	Sequence 258, App
17	445.5	37.7	87563	3	US-09-453-702B-57	Sequence 57, Appl
18	431.5	36.5	4403765	3	US-09-103-840A-2	Sequence 2, Appli
19	431.5	36.5	4411529	3	US-09-103-840A-1	Sequence 1, Appli
20	300.5	25.4	462	4	US-09-543-681A-1913	Sequence 1913, Ap
21	281.5	23.8	312	4	US-09-489-039A-2090	Sequence 2090, Ap
22	267.5	22.7	300	5	PCT-US96-05320A-543	Sequence 543, App
23	266.5	22.6	324	4	US-09-543-681A-1929	Sequence 1929, Ap
24	264.5	22.4	399	3	US-09-134-001C-2212	Sequence 2212, Ap
25	261	22.1	438	3	US-09-134-001C-2145	Sequence 2145, Ap
26	261	22.1	609	4	US-09-602-777A-9	Sequence 9, Appli
27	261	22.1	3164	4	US-09-710-279-3828	Sequence 3828, Ap
28	261	22.1	3234	4	US-09-710-279-3853	Sequence 3853, Ap
29	255.5	21.6	312	4	US-09-328-352-1761	Sequence 1761, Ap
30	250.5	21.2	378	4	US-09-252-991A-10313	Sequence 10313, A
31	250.5	21.2	423	4	US-09-602-777A-17	Sequence 17, Appl
32	250.5	21.2	813	4	US-09-252-991A-10547	Sequence 10547, A
33	250.5	21.2	831	4	US-09-252-991A-10512	Sequence 10512, A
34	247	20.9	336	4	US-09-710-279-115	Sequence 115, App
35	239.5	20.3	360	4	US-09-328-352-1789	Sequence 1789, Ap
36	230.5	19.5	306	4	US-09-601-198-92	Sequence 92, Appl
37	218	18.5	414	4	US-09-252-991A-10315	Sequence 10315, A
38	215.5	18.2	375	4	US-09-601-198-91	Sequence 91, Appl
39	208	17.6	330	4	US-09-489-039A-2095	Sequence 2095, Ap
40	119	10.1	789	4	US-09-252-991A-10654	Sequence 10654, A
41	117	9.9	220	4	US-09-602-777A-11	Sequence 11, Appl
42	117	9.9	2259	4	US-09-710-279-4078	Sequence 4078, Ap
43	94	8.0	1664976	4	US-08-966-521B-1	Sequence 1, Appli
44	94	8.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
45	92	7.8	870	3	US-09-134-001C-1578	Sequence 1578, Ap

## ALIGNMENTS

RESULT 1  
US-08-920-095-1  
; Sequence 1, Application US/08920095  
; Patent No. 5837240  
; GENERAL INFORMATION:  
; APPLICANT: Cynthia K. Lee et al.  
; TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/920,095  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/431,041  
; FILING DATE: 28-APR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 2735 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-920-095-1

## Alignment Scores:

Pred. No.:	1-2e-70	length:	2735
Score:	610.50	Matches:	118
Percent Similarity:	71.88%	Conservative:	43
Best Local Similarity:	52.68%	Mismatches:	62
Query Match:	51.69%	Indels:	1
DB:	2	Gaps:	1

US-09-904-994B-2 (1-226) x US-08-920-095-1 (1-2735)

QY 1 VallyLeuThrProlysglucingluysPheLeuLeuTyrrYrAlaglygluVala 20  
DB 102 ATGAACTCACCCCAAAAGATTAGATTGATGCTCCTACCTGAGATTGGCT 161  
QY 21 ArglyValgylsAlaglyleuylsLeuAerGlnProgluAllealatyrlleser 40  
DB 162 AAAAAAGCAAGAAAGGCAATTAACTTAACTATGTAAGACAGTAGCTTGAATTAG 221  
QY 41 AlaHieIlleMerAerGlnuAaArgsglylylVleVleThrValAlaglnleuMetglu 60  
DB 222 GCCCATATTATGAAAG 281  
QY 61 CysMetHiePheLeuylsAerGlnuValMetProglYValGlyAsnMetValProAer 80  
DB 282 GGGCGGCACTCTTTTAAACAGATGATGTGATGATGGCGTGCAAGATGATCCATGAA 341  
QY 81 LeuGlyValGluAlaThrPheProAerGlyThrlyLeuValThrValAsnTrpProIle 100  
DB 342 GTGGGTATTGAAGCGATGTTTCTGTATGGAGACTAACTCGTAACCGTGCATCCCTATT 401  
QY 101 GluProAerGlnHiePheLySAlaglygluVallyPheGlyCysAerlyAerIleGlu 120  
DB 402 GAGGCCAATGTAATTAATGTTCTGCTGAGTGTG---TTCTTAAAAAATGAAAGCATCACT 458  
QY 121 LeuAsnAlaglylysgluValThrGluLeuGluValThrAsnGluGlyProlySerLeu 140  
DB 459 ATCAACAGAGGCAAAAGAGCGGTTAGCGTGAAGATTAAAGTTGTGGCGCAGACCGGTT 518  
QY 141 HieValGlySerHiePheHiePhePheGlnuAlaAsnlyValleuylsPheAerGln 160  
DB 519 CAATGCGCTCACTTCATTTCTTTGAAGTGAATGATGAGCTTGAACATTTGACAGAGAA 578  
QY 161 lySAlaTyrrGlyAerGlnuAerIleProSerGlyAsnThrleuAerIleGlyAlaGly 180  
DB 579 AAAAATTCGCTAAACGCTTAAGCATTTGCGAGCGGAGACCGGTAAGATTGAGCTTGGC 638  
QY 181 GlnThrArglyValGlnleuIleProLeuGlyGlySerlyVallyleGlyMetAer 200  
DB 639 GAAAGAAAATCCCTGAATGATTGACATTTGCGGTAACAGAGAATCTTTGATTTAAC 698  
QY 201 GlyLeuValAsnAsnIleAlaAerGlnuArgHieLyHieValleuAerlySAlaLy 220  
DB 699 GCATTGTGATGACAGACAGACAGACAGACAGACAGACAGACAGACAGACAGATAA 758  
QY 221 SerHieGlyPhe 224  
DB 759 GAGCGTGGTTTT 770

## RESULT 2

PCT-US96-05800-1

Sequence 1, Application PC/TUS9605800

GENERAL INFORMATION:

APPLICANT: Oryax, Inc.

TITLE OF INVENTION: MULTIMERIC, RECOMBINANT UREASE VACCINE

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish &amp; Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/05800

FILING DATE: 23-APR-1996

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/431,041

FILING DATE: 28-APR-1995

CLASSIFICATION:

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: 08/568,122

FILING DATE: 06-DEC-1995

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Clark, Paul T.

REGISTRATION NUMBER: 30,162

REFERENCE/DOCKET NUMBER: 06132/020001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 542-5070

TELEFAX: (617) 542-8906

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2735 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

PCT-US96-05800-1

## Alignment Scores:

Pred. No.:	1-2e-70	length:	2735
Score:	610.50	Matches:	118
Percent Similarity:	71.88%	Conservative:	43
Best Local Similarity:	52.68%	Mismatches:	62
Query Match:	51.69%	Indels:	1
DB:	5	Gaps:	1

US-09-904-994B-2 (1-226) x PCT-US96-05800-1 (1-2735)

QY 1 VallyLeuThrProlysglucingluysPheLeuLeuTyrrYrAlaglygluVala 20  
DB 102 ATGAACTCACCCCAAAAGATTAGATTGATGCTCCTACCTGAGATTGGCT 161  
QY 21 ArglyValgylsAlaglyleuylsLeuAerGlnProgluAllealatyrlleser 40  
DB 162 AAAAAAGCAAGAAAGGCAATTAACTTAACTATGTAAGACAGTAGCTTGAATTAG 221  
QY 41 AlaHieIlleMerAerGlnuAaArgsglylylVleVleThrValAlaglnleuMetglu 60  
DB 222 GCCCATATTATGAAAG 281  
QY 61 CysMetHiePheLeuylsAerGlnuValMetProglYValGlyAsnMetValProAer 80  
DB 282 GGGCGGCACTCTTTTAAACAGATGATGTGATGATGGCGTGCAAGATGATCCATGAA 341  
QY 81 LeuGlyValGluAlaThrPheProAerGlyThrlyLeuValThrValAsnTrpProIle 100  
DB 342 GTGGGTATTGAAGCGATGTTTCTGTATGGAGACTAACTCGTAACCGTGCATCCCTATT 401  
QY 101 GluProAerGlnHiePheLySAlaglygluVallyPheGlyCysAerlyAerIleGlu 120  
DB 402 GAGGCCAATGTAATTAATGTTCTGCTGAGTGTG---TTCTTAAAAAATGAAAGCATCACT 458



```
APPLICANT: Londono-Arcila, Patricia
APPLICANT: Freeman, Donna
TITLE OF INVENTION: Use of salmonella vectors for
FILE REFERENCE: 06132/060001
CURRENT APPLICATION NUMBER: US/09/431,705
CURRENT FILING DATE: 1999-11-01
NUMBER OF SEQ ID NOS: 52
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 19
LENGTH: 4824
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: includes sequences from Helicobacter pylori,
OTHER INFORMATION: Salmonella typhimurium, and Escherichia coli
NAME/KEY: CDS
LOCATION: (3893)...(3934)
NAME/KEY: CDS
LOCATION: (3938)...(4027)
NAME/KEY: CDS
LOCATION: (4031)...(4285)
NAME/KEY: CDS
LOCATION: (4289)...(4300)
NAME/KEY: CDS
LOCATION: (4304)...(4408)
NAME/KEY: CDS
LOCATION: (4412)...(4471)
NAME/KEY: CDS
LOCATION: (4475)...(4588)
NAME/KEY: CDS
LOCATION: (4592)...(4669)
NAME/KEY: CDS
LOCATION: (4673)...(4711)
NAME/KEY: CDS
LOCATION: (4715)...(4774)
NAME/KEY: CDS
LOCATION: (4784)...(4824)
US-09-431-705-19
```

```
Alignment Scores:
Pred. No.: 2.9e-70 Length: 4824
Score: 610.50 Matches: 118
Percent Similarity: 71.88 Conservative: 43
Best Local Similarity: 52.68 Mismatches: 62
Query Match: 51.69 Indels: 1
DB: Gaps: 1
```

US-09-904-994B-2 (1-226) x US-09-431-705-19 (1-4824)

```
QY 1 VallyleuThProlysgIngulvPheleuTyTrrAlaGlyValAla 20
DB 86 ATGAACCTCAACCCAAAGAGTTAGATGATGCTCCACATCCCTGAGATTGGCT 145
QY 21 ArgLySArgLyLeuAlaGlyLeuLysLeuAenGlnProGluAlaIleAlaTyrIleSer 40
DB 146 AAAAAACGCAAGAAAGCAATTAAGCTTAATAGTGAAGACATAGCTTGAATAGT 205
QY 41 AlaHisIleMetAspGluAlaArgArgGlyLysLysThrValAlaGlnLeuMetGluGlu 60
DB 206 GCCCATATATATGAGAAAGACGAGAGCTGGTAAAGAAAGCTGCCGCTGAATTGATGCAAGAA 265
QY 61 CysMetHisPheLeuLysLysAspGluValIleMetProGluValGlyLysMetValProAsp 80
DB 266 GGGCGGCACTCTTTTAAACCAAGATGATGATGATGGCGTGGCAAGCATGATCCATGAA 325
QY 81 LeuGlyValGlyValAlaThrPheProAspGlyThrLysLeuValThrValAsnTyrProIle 100
DB 326 GTGGGATATGAAGCGAGATGTTTCTGATGGACATAAAGCTGTAACCGTGCATACCCCTATT 385
QY 101 GluProAspGluHisPheLysAlaGlyGlyValLysPheGlyCysAspLysAspIleGlu 120
DB 386 GAGGCGCAATGATTAATTAGTTCTGATGAGTGTG---TTCTTAAATAATGAAGACATCACT 442
```

```
QY 121 LeuAsnAlaGlyLysGlyValThrGluLeuGluValThrAsnGluGlyProLysSerLeu 140
DB 443 ATCAACGAAAGCAAAAGAACCCCTTAGCTGGAAGTTAAAGTTGGCAGACAGCCGTT 502
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
DB 503 CAATCGGCTCACACTTCATTTCTTTGAAAGTAATGATGCTTACATTTGACAGAGAA 562
QY 161 LysAlaIleGlyLysAlaGlyLeuAspIleProSerGlyAsnThrLeuAlaGlyIleGlyAlaGly 180
DB 563 AAACTTCGTGTAACCGCTTACATTCGACATTCGACGCGGACAGCGGTAAGTTAGCCTGSC 622
QY 181 GlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200
DB 623 GAAGAAAATCCGTGAATGATTGACATTCGCGGTAAAGAGAATCTTTGGATTATAC 682
QY 201 GlyLeuValAsnAsnIleAlaAspGluArgHisLysAlaLeuAspLysAlaLys 220
DB 683 GCATTGGTTGATAGACAAGCAGACAGAACCAAAATTCCTTACACAGAGCTTAA 742
QY 221 SerHisGlyPhe 224
DB 743 GAGCGTGTTT 754

RESULT 5
US-08-467-822-19
Sequence 19, Application US/08467822
Patent No. 5843460
GENERAL INFORMATION:
APPLICANT: Labigne, Agnes
APPLICANT: Sauerbaum, Sebastian
APPLICANT: Ferrero, Richard L.
APPLICANT: Thiberge, Jean-Michel
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flanagan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,822
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/447,177
FILING DATE: 19-MAY-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/432,697
FILING DATE: 02-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495, 0137-02000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
```



```

; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name="Shine-Dalgarno
; OTHER INFORMATION: sequence."
;
; US-08-467-822-19
;
Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
DB: Gaps: 1

US-09-904-994B-2 (1-226) x US-08-467-822-19 (1-2619)
Qy 1 VallyleuThrProlysgIngluLysePheleuTyrrYalaglYluVala 20
Db 43 ATGAACTAAGCGCTAAAGAACTAGACAGATTATGCTTATCGGGGAGATTGCA 102
Qy 21 ArglyeArglyslaglYleuLyseleuInProgluAlaileAlaTyrlIser 40
Db 103 GAAGAACGCTGGCGCTGGTGTGAACATTAACACGAGCGGCTCATTAAGC 162
Qy 41 AlAlleleMerAaspGluAlaArgArglylyseThrValaglInleuMetGlu 60
Db 163 GGGCGTGTGATGAAAGGCGCGGATGTAATAAGCGGCGGATTTGATGCAAGAA 222
Qy 61 CyMetHisPheleuLysePheGluValMetProglYValGlyAsnMetValProasp 80
Db 223 GCGAGGACTTGGCTTAAAGAAAGAAATGTATGAGCGGCTAGACACATGATTCAGAA 282
Qy 81 LeuGlYValGluAlaThrPheProaspGlyThrlYleuValThrValAsnTTPProile 100
Db 283 GTGGGATTAAGCTAACTTCCCGAGTGAACCAAGCTTGAATCACTCCGCGTA 342
Qy 101 GluProaspGluHisPheLyslaglYValLysePheGlyCyAspLyAspIleGlu 120
Db 343 GAGGATTAATGCAAAATTAGCCCCGCGAGGATC---TTCTTAAAAATAGAGACATTACT 399
Qy 121 LeuAsnAlaglYlyseGluValThrGluLeuGluValThrAsnGluGlyProlySerleu 140
Db 400 ATTAACGCGCGCAAGAAAGCATTAAGTGAATAAATAAAGCGATTCGCTGTG 459
Qy 141 HisValGlySerHisPheHisPhePheGluAlaAsnLyslaleuLyPheAspArgGlu 160
Db 460 CAGGTGGATCAACATTTCCACTTTCGAGTGAATTAAGCTTGGACCTTCGATCGGCA 519
Qy 161 LyAlaTyrglylyslYleuAspIleProSerGlyAsnThrleuValGileGlyAlaglY 180
Db 520 AAAAGCTTTGCAAAAGCCCTTAGACATTGCTGGAACAGCGGCGCTTGAACCGGG 579
Qy 181 GluThrArglyslYslInleuIleProleuGlylyseIlyslYslValIleGlyMetAsn 200
Db 580 GAGGAAAAAGTGTGGAATCATTTGACATCGCGGGAATAGCGCATCTATGCTTAAT 639
Qy 201 GlyleuValAsnAsnIleAlaAspGluArgHislyHislyslaleuAspLyslAlaLy 220
Db 640 TCTTGTGGATGCGCAAGCGCATGCGATGTAAAAAAGCTTGAACGCGCTTAA 699
Qy 221 SerHisGlyPhe 224
|||
```

```

Db 700 GAAAGGTTTT 711
RESULT 6
US-08-432-697-19
; Sequence 19, Application US/08432697
; Patent No. 6248330
; GENERAL INFORMATION:
; APPLICANT: Labigne, Agnes
; APPLICANT: Sauerbaum, Sebastien
; APPLICANT: Ferreira, Richard L.
; APPLICANT: Thibierge, Jean-Michel
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
; TITLE OF INVENTION: HELICOBLACTER INFECTION, POLYPEPTIDES FOR USE IN THE
; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSER: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0137-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard_name="Shine-Dalgarno
; OTHER INFORMATION: sequence."
;
; US-08-432-697-19
;
Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
DB: Gaps: 1

US-09-904-994B-2 (1-226) x US-08-432-697-19 (1-2619)
Qy 1 VallyleuThrProlysgIngluLysePheleuTyrrYalaglYluVala 20
Db 43 ATGAACTAAGCGCTAAAGAACTAGACAGATTATGCTTATCGGGGAGATTGCA 102
```

```

QY 21 ArglysArglyAlaGluGlyLeuAsnGlnProGluAlaIleAlaTyrIleSer 40
DB 103 GAAGAACGCTGGCGGCTGTGTGAATCTCAATTACGACGAGCGGCTCCGCTTAC 162
QY 41 AlaHisIleMetAspGluAlaArgArgGlyLysLysThrValAlaGlnLeuMetGluGlu 60
DB 163 GCGCGGTGATGAGAAAGCGCGGTGATGATTAATAAAGCGTGGCGGATTTGATGCAAGA 222
QY 61 CysMetHisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAsp 80
DB 223 GCGAGACCTGGCTTAAATAAAGAAATGTGATGACGCGGTGACCAAGCATGATTCATGA 282
QY 81 LeuGlyValGluAlaThrPheProAspGlyThrLysLeuValThrValAsnTrpProIle 100
DB 283 GTGGGATTTGAAGCTAATCTCCCGCATGACCAAGCATTTGTAATTCACACTCCGGTA 342
QY 101 GluProAspGluHisPheLysAlaGlyLysValLysPheGlyCysAspLysAspIleGlu 120
DB 343 GAGGATTAATGCAAAATTAGCCCGCGGAGGTC---TTCTTAAATAATGAGACATTACT 399
QY 121 LeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLysSerLeu 140
DB 400 ATTAACGCCCGCAAAAGCCATTAGCTTGAAGTAAATAAATGAAGCGATGCTCTGTG 459
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
DB 460 CAGGTGGGATTCATTTCCACTTCTTGCAGAGTAAATACCTCTTGACCTTCATCCGCGA 519
QY 161 LysAlaTyrGlyLysArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
DB 520 AAAAGCTTTGCAAGCGCTAGCATGTGATGATGCAAGCGGTGCGCTTGAACCGCGG 579
QY 181 GlnThrArgLysValGlnLeuLysProLeuGlyGlySerLysLysValIleGlyMetAsn 200
DB 580 GAGGAAAAAGTGTGACTCATTCATGACATCGCGGAAATGAAGGCAATCATGAGCTTTAAT 639
QY 201 GlyLeuValAsnAsnIleAlaAspGluArgHisLysHisLysAlaLeuAspLysAlaLys 220
DB 640 TCTTGTGTGATGCCCAAGCCGATGCCGATGGTAAATAAATCGCGCTTAAAGCGCTAA 699
QY 221 SerHisGlyPhe 224
DB 700 GAAAAAGGTTTT 711

```

RESULT 7  
 US-08-466-248-19  
 ; Sequence 19, Application US/08466248  
 ; Patent No. 6258359  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Labigne, Agnes  
 ; APPLICANT: Sauerdam, Sebastien  
 ; APPLICANT: Ferrero, Richard L.  
 ; APPLICANT: Thiberge, Jean-Michel  
 ; TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
 ; TITLE OF INVENTION: HELICOCTER INFECTION, POLYPEPTIDES FOR USE IN THE  
 ; TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
 ; NUMBER OF SEQUENCES: 44  
 ; CORRESPONDENCE ADDRESSES:  
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
 ; STREET: 1300 I Street, N.W.  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: USA  
 ; ZIP: 20005-3315  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/466,248

```

; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/447,177
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,697
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495, 0137-02000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2619 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 31..36
; OTHER INFORMATION: /standard name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence.
; NAME/KEY: misc feature
; LOCATION: 756..759
; OTHER INFORMATION: /standard name= "Shine-Dalgarno"
; OTHER INFORMATION: sequence.
US-08-466-248-19

Alignment Scores:
Pred. No.: 1,3e-65 Length: 2619
Score: 572.50 Matches: 113
Percent Similarity: 69.64% Conservative: 43
Best Local Similarity: 50.45% Mismatches: 67
Query Match: 48.48% Indels: 1
Gaps: 1
DB: 3

US-09-904-994B-2 (1-226) x US-08-466-248-19 (1-2619)
QY 1 ValLysLeuThrProLysGluGlnGlyLysPheLeuLeuTyrTyrAlaGlyGluValAla 20
DB 43 ATGAATACTTAACCCCTTAAGAACTAGACAAATTAACTCCATTATGCGGCGAGATTGGCA 102
QY 21 ArglysArglyAlaGluGlyLeuAsnGlnProGluAlaIleAlaTyrIleSer 40
DB 103 GAAGAACGCTGGCGGCTGTGTGAATCTCAATTACGACGAGCGGCTCCGCTTAC 162
QY 41 AlaHisIleMetAspGluAlaArgArgGlyLysLysThrValAlaGlnLeuMetGluGlu 60
DB 163 GCGCGGTGATGAGAAAGCGCGGTGATGATTAATAAAGCGTGGCGGATTTGATGCAAGA 222
QY 61 CysMetHisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAsp 80
DB 223 GCGAGACCTGGCTTAAATAAAGAAATGTGATGACGCGGTGACCAAGCATGATTCATGA 282
QY 81 LeuGlyValGluAlaThrPheProAspGlyThrLysLeuValThrValAsnTrpProIle 100
DB 283 GTGGGATTTGAAGCTAATCTCCCGCATGACCAAGCATTTGTAATTCACACTCCGGTA 342
QY 101 GluProAspGluHisPheLysAlaGlyLysValLysPheGlyCysAspLysAspIleGlu 120
DB 343 GAGGATTAATGCAAAATTAGCCCGCGGAGGTC---TTCTTAAATAATGAGACATTACT 399
QY 121 LeuAsnAlaGlyLysGluValThrGluLeuGluValThrAsnGluGlyProLysSerLeu 140
DB 400 ATTAACGCCCGCAAAAGCCATTAGCTTGAAGTAAATAAATGAAGCGATGCTCTGTG 459

```

QY 141 HisValGlySerHisPheHisPheGluAlaAsnValAlaLeuValPheAspArgGlu 160  
Db 460 CAGGTGGATGACATTTCCACTTCTTCGAGTGAATGAGCTTGGACCTGATCGCCCA 519  
QY 161 LysAlaTyrGlyLysArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyValAlaGly 180  
Db 520 AAAAGCTTTTCCAAAGCCCTTGAACATTCATTCGAAACAGCGGTCGCTTGAACCCGG 579  
QY 181 GluThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200  
Db 580 GAGAAAAAAGTGTGGAATCATTCATTCGCGCGGGAATTAAGCGATCATGGCTTAAAT 639  
QY 201 GlyLeuValAlaAsnIleAlaAspGluArgHisLysHisLysValAlaLeuAspLysVal 220  
Db 640 TCTTGTGGATGCCCAAGCCGATGCCGATGTAAATACTCGCTTAAACGGGCTTAA 699  
QY 221 SerHisGlyPhe 224  
Db 700 GAAAAAGCTTT 711  
RESULT 8  
US-09-710-279-4032  
; Sequence 4032, Application US/09710279  
; Patent No. 6703492  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/09/710,279  
; CURRENT FILING DATE: 2000-11-09  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4032  
; LENGTH: 3475  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-09-710-279-4032  
Alignment Scores:  
Pred. No.: 4,96e-55 Length: 3475  
Score: 494.50 Matches: 109  
Percent Similarity: 64.41% Conservative: 34  
Best Local Similarity: 49.10% Mismatches: 63  
Query Match: 41.87% Indels: 16  
Gaps: 4  
US-09-904-994b-2 (1-226) x US-09-710-279-4032 (1-3475)  
QY 4 ThrProLysGlnGlnGlnLysPheLeuLeuTyrTrpAlaGlyGluValAlaArgLysArg 23  
Db 1864 ACACAAAGCTGAACAGCAAAATTGATATGTTAGCTGTGAGGTGACGTCGAGGA 1923  
QY 24 LysAlaGlnGlyLeuLysLeuAsnGlnProGlnAlaIleAlaTyrIleSerAlaHisIle 43  
Db 1924 AAAGCAAGAGGACTTAATCACTTCAAGACACTGCTTAAATCAGAGAGGAATTA 1983  
QY 44 MetAspGluAlaArgArgGlyLysLysThrValAlaGlnLeuMetGlnGluCysMetHis 63  
Db 1984 TTAGAAAGCCGCGCTGATGCT---AAAACGCTAGCTGACATGAGCTATGAAAAACA 2040  
QY 64 PheLeuLysLysAspGluValMetProGlnValAlaGlnMetValProAspLeuGlyVal 83  
Db 2041 ATTTTAAACGAGAGATGTCATGATGAGTGGCTAGCTTAACTGATTAACAGAACTTGA 2100  
QY 84 GluAlaThrPheProAspGlyThrLysLeuValThrValAsnTyrProIle----- 100  
Db 2101 GAAGCAACTTTTCCAGATGGTACTAAGTTTATAACAGTCCATCCCAATGCTTTAAGGA 2160

QY 101 -----Glu-ProAspGluHisPheLysAlaGlyGlnValLysPheGlyCysAs 116  
Db 2161 GCGTATATCAATGATGCTTGGTAAATTAAT------GTAAA-----AA 2199  
QY 116 PheAspIleGlnLeuAsnAlaGlyLysGlnValThrGluLeuGlnValThrAsnGln 136  
Db 2200 TACTGAAATGAACTCAATTAACATCATCCGAAACGGTGTATTAAGTAAAAATACAG 2259  
QY 136 YProLysSerLeuHisValGlySerHisPheHisPheGluAlaAsnValAlaLeu 156  
Db 2260 CGATAGACCTATACAAATGATGTTCACTTTTCCATTTTTCGAAACCAATAAAGCATTA 2319  
QY 156 PheAspArgGlnLysValAlaTyrGlyLysArgLeuAspIleProSerGlyAsnThrLeu 176  
Db 2320 ATTTGATGCTGAGAAACATATGTTAAACATTTGATATTTCTCGAGAGCTGACGTGAG 2379  
QY 176 GLeGlyValGlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysVal 196  
Db 2380 ATTTGAACTGGAGATGAAAAAAGTACACTTGTGAATATTCTGACGACGATAAAT 2439  
QY 196 LLeGlyMetAsnGlyLeuValAlaAsnIleAlaAspGluArgHisLysHisLysVal 216  
Db 2440 TTATGATTCGCTGCTTATGATGATGCGATGATGAGAAAGCGGATTTCCGTCCAA 2499  
QY 216 uAsp 217  
Db 2500 TGAT 2503  
RESULT 9  
US-08-956-171E-22/C  
; Sequence 22, Application US/08956171E  
; Patent No. 6593114  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
Gil H. Choi  
Patrick S. Dillon  
Craig A. Rosen  
Steven C. Barash  
Michael R. Fannon  
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 5256  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: Human Genome Sciences, Inc.  
STREET: 9410 Key West Avenue  
CITY: Rockville  
STATE: Maryland  
COUNTRY: USA  
ZIP: 20850  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage  
COMPUTER: HP Vectra 486/33  
OPERATING SYSTEM: MSDOS version 6.2  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/956,171E  
FILING DATE: 20-Oct-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/009,861  
FILING DATE: January 5, 1996  
APPLICATION NUMBER: 08/781,986  
FILING DATE: January 3, 1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Mark J. Hyman  
REGISTRATION NUMBER: 46,789  
REFERENCE/DOCKET NUMBER: PB248P1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (240) 314-1224  
TELEFAX: (301) 309-8439  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 596 base pairs





```

1 STREET: 2190 Commons Parkway
2 City: Okemos
3 STATE: Michigan
4 COUNTRY: USA
5 ZIP: 48864
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb
9 MEDIUM TYPE: storage
10 COMPUTER: IBM Compatible
11 OPERATING SYSTEM: MS-DOS
12 SOFTWARE: Wordperfect 5.1
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/687,645B
16 FILING DATE: July 26, 1996
17
18 CLASSIFICATION: 435
19
20 PRIOR APPLICATION DATA:
21 APPLICATION NUMBER:
22 FILING DATE:
23
24 ATTORNEY/AGENT INFORMATION:
25 NAME: Ian C. McLeod
26
27 REGISTRATION NUMBER: 20,931
28 REFERENCE/DOCKET NUMBER: MSU 4.1-309
29
30 TELECOMMUNICATION INFORMATION:
31 TELEPHONE: (517) 347-4100
32 TELEFAX: (517) 347-4103
33
34 TELEX: No. 5846752e
35
36 INFORMATION FOR SEQ ID NO: 1:
37 SEQUENCE CHARACTERISTICS:
38 LENGTH: 2400
39 TYPE: nucleotides
40 STRANDEDNESS: Single
41 TOPOLOGY: Linear
42
43 MOLECULE TYPE:
44 DESCRIPTION: cDNA
45
46 HYPOTHEICAL: NO
47
48 ANTI-SENSE: NO
49
50 ORIGINAL SOURCE:
51 ORGANISM: Klebsiella aerogenes
52 STRAIN: CG253
53
54 INDIVIDUAL ISOLATE:
55 CELL TYPE: N/A
56
57 FEATURE:
58 NAME/KEY: cDNA encoding mutant urease
59 NAME/KEY: 'H2190
60 LOCATION: Modification at position 1312 to
61 LOCATION: glutamine
62 IDENTIFICATION METHOD: Sequencing
63 OTHER INFORMATION:
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
```

```

Db      208 GCGCGTCACGTCCTCTGACCCGCGACGAGTGAGTGGGCGTCCGGAATAATATCCGGAT 267
Oy      81 LeuGIyValIgluAlaThrPheProAspGIyThrIyLseuValIThrValAentIrpProIle 100
Db      268 ATCCAGGTGGAAGCCACTTCCCGGACGGCTCGAAGCTGTGCACCGTTCCACACCCGATT 327
Oy      101 -----GluProAspGIu---HisPheIyAspIagIyGIuValIyPhe 113
Db      328 ATCTGAGGTAGCGCCCATGATCCCGGTGAATATCCGTTAAGCCCGGTCA----- 378
Oy      114 GIYcAspIyAspIleGIeIuleuAsnIaGIyIySGIuValIThrGIuLeuGIuValIThr 133
Db      379 -----ATAGCCCTGAATACCGCGCGCAACCTGTCCGTGCTGTGAG 423
Oy      134 AsnGIuGIyProIySerIeuHisValGIySerHisPheHisPhePheGIuAlaAsnIyS 153
Db      424 AACCAACGCGCATGCGCGCATTCAGGTGCGGTTCGCACTACATTCGCGCAGGTAAACCG 483
Oy      154 AlAlenIyPheAspARGIuIyValIyValIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 173
Db      484 GCGCTGAAGTTCACCGTCAGACGAGCGCGGTATTCCTGAATATTCGCGCGGACG 543
Oy      174 ThrIeuARGIleGIyAlaGIyIyInThrIyIyValIyIyIyIyIyIyIyIyIyIyIy 193
Db      544 GCGGTACGCTTGAACCCGCGCCAGAAACGACGAGCTGTGTGCGCTTGCCGCTCAC 603
Oy      194 IyAsIyValIleGIyMetAsnGIyIeuVal 203
Db      604 CGCGCGTCTTCGCGCTTCGCGCGGAGGTG 633

RESULT 13
US-07-732-242C-8
; Sequence 8, Application US/07732242C
; Patent No. 5298399
; GENERAL INFORMATION:
; APPLICANT: Uozumi, Takeehi; Maesaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yoneta, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frieauf, Holtz, Goodman & Woodward, P. C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII Form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET NUMBER: 910532/HG
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)972-1400
; TELEFAX: (212)370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6131 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: both

```

; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; US-07-732-242C-8  
 Alignment Scores:  
 Pred. No.: 3,486-51 Length: 6131  
 Score: 468.50 Matches: 101  
 Percent Similarity: 59.59% Conservative: 45  
 Best Local Similarity: 41.22% Mismatches: 71  
 Query Match: 39.67% Indels: 28  
 DB: 1 Gaps: 4  
 US-09-904-994B-2 (1-226) x US-07-732-242C-8 (1-6131)  
 QY 1 VallyLeuThrProlyseGluGluLeuPheLeuLeuTyrrYrAlaGlyGluValAla 20  
 Db ATGAACTGCTTCACGTGAAGTGAAGAAAGCTCATGATGAGTGGCGGCTGACTGGCC 397  
 QY 21 ArgGlyAspGlyValAlaGluGlyLeuLeuAsnGlnProGluAlaIleAlaTyrIleSer 40  
 Db CGCGCTGTAAGAGCGGGGCTTAATAATTATCTGAAGCTGCTGCAATGATTAACA 457  
 QY 41 AlaHisIleMetAspGluAlaArgArgGlyLeuSerThrValAlaGlnLeuMetGluGlu 60  
 Db TATGAAGTGTCTGAGGGGGCGGGGATGCA---AAACGCTAGCTCACTTAATCAATAC 514  
 QY 61 CysMetHisPheLeuLeuValysAspGluValMetProGlyValGlyAsnMetValProAsp 80  
 Db 515 GGTGCAACGATCTTACAAAAGATGTATGAGAGGGGTGGCCGAAATGATCCGGAT 574  
 QY 81 LeuGlyValGluAlaThrPheProAspGlyThrIleLeuValThrValAsnTrpProIle 100  
 Db 575 ATTCAAAATGAGGCACTTTCCTGATGAAACAAAGCTTGACGCTTCATGACCCGATC 634  
 QY 101 -----Glu-ProAspGlnHisPheLeuValysGluValVally 112  
 Db 635 CGTTAAATGAGAGAGAGTACGATGATACAGGGGAGATGTATTAATAAAGAA---- 689  
 QY 112 sPheGlyCysAspGlyAspIleGlyLeuAsnAlaGlyLeuGluValThrGluLeuGluVal 132  
 Db 690 -----CCTATTATGCAATCAAAATGACAGACGATCAAGATTCGCGT 733  
 QY 132 LThrAsnGluGlyProlyseSerLeuHisValGlySerHisPheHisPheGluAlaAs 152  
 Db 734 GTTAAACCGGGCGATGACCTGTCAGGTGGTTCCTCATTTTATTTTGAAGTGA 793  
 QY 152 nlyValAlaLeuLysPheAspArgGlyValArgGlyValysArgLeuAspIleProSerG 172  
 Db 794 TCATCGCTTCATTTTCATGCGTGAAGAAAGCATTTGGCGCTTGAAATATTCGCGCTG 853  
 QY 172 yAsnThrLeuAspGlyIleGlyAlaGlyIleThrArgLeuValGlnLeuIleProLeuGlyG 192  
 Db 854 AACCGCGCTTCGCTTCAGCCCGGAGATGCGAAAGAAAGTAAATTAATTCATTTTCA 913  
 QY 192 ySerLeuValValIleGlyMetAsnGlyLeuValAsnAsnIleAlaAsp----- 208  
 Db 914 TGAACGCAAGTGTATGATGTTAAATTAATTAACCAATGATCAGTTGAATGGGAAAG 973  
 QY 209 -----GluArgHisLeuHisValAlaLeuAspGlyValys 221  
 Db 974 AAATAGATTTTCGATGTCTCGAAAGCAATATGCGGATATGTTTGAACCAAGTGGGC 1033  
 QY 221 rHisGlyPheIle 225  
 Db 1034 GACCGCATTCGTT 1046  
 RESULT 14  
 ; US-09-557-884-1/c  
 ; Sequence 1, Application US/09557884  
 ; Patent No. 6506581  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fleischmann et al.  
 ; TITLE OF INVENTION: The Nucleotide sequence of

; the Haemophilus influenzae Rd Genome, Fragments  
 ; thereof, and Uses Thereof  
 ; NUMBER OF SEQUENCES: 1  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Human Genome Sciences, Inc.  
 ; STREET: 9410 Key West Avenue  
 ; CITY: Rockville  
 ; STATE: MD  
 ; COUNTRY: USA  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3 1/2 inch diskette  
 ; COMPUTER: Dell Pentium  
 ; OPERATING SYSTEM: MS DOS v6.22  
 ; SOFTWARE: ASCII Text  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/557,884  
 ; FILING DATE: 25-Apr-2000  
 ; CLASSIFICATION: <Unknown>  
 ; PRIORITY APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/476,102  
 ; FILING DATE: JUN-5-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Michelle S. Marks  
 ; REGISTRATION NUMBER: 41,971  
 ; REFERENCE/DOCKET NUMBER: PB186P3  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 301-309-8504  
 ; TELEFAX: 301-309-8439  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1830121 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ; US-09-557-884-1  
 Alignment Scores:  
 Pred. No.: 3,796-47 Length: 1830121  
 Score: 467.00 Matches: 96  
 Percent Similarity: 61.61% Conservative: 42  
 Best Local Similarity: 42.86% Mismatches: 62  
 Query Match: 39.54% Indels: 24  
 DB: 4 Gaps: 3  
 US-09-904-994B-2 (1-226) x US-09-557-884-1 (1-1830121)  
 QY 3 LeuThrProlyseGluGluLeuPheLeuLeuTyrrYrAlaGlyGluValAlaArgLys 22  
 Db 56486 TTAACCTTCAGAGAAACAAAGAAAGCTGATCTTTCCTCGCGGCGGAATCTGGCAAAA 564809  
 QY 23 ArgGlyValGluGlyLeuLeuAsnGlnProGluAlaIleAlaTyrIleSerAlaHis 42  
 Db 564808 CGCAACACAGCGCGGCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 564749  
 QY 43 IleMetAspGluAlaArgArgGlyLeuSerThrValAlaGlnLeuMetGluGluCysMet 62  
 Db 564748 TTACAAAGCAGCAGAAAGAAAGG---ATGAGTGTAGCGGAAGTCATGATATGGGCA 564692  
 QY 63 HisPheLeuLysAspGlyValMetProGlyValGlyAsnMetValProAspLeuGly 82  
 Db 564691 ACACTTTTAACCGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 564632  
 QY 83 ValGluAlaThrPheProAspGlyThrLeuLeuValThrValAsnTrpProIleGlu--- 101  
 Db 564631 ATTGAAGTACTTTCCTCCGATGGCAGCAAACTTGTACCTGATATCAATCAGATAA 564572  
 QY 101 ----- 101  
 Db 564571 CCGWAGGTGGGCTTTAGCCCAAAATAAATATCATATGATGCTGAAGCCCAACCT 564512  
 QY 102 ---ProAspGlnHisPheLeuValysGluValValysPheGlyCysAspGlyAspIleGlu 120

```

Db 564511 ACAAGAACAAATATGATCCAGCGCAATACCAATTA---GCCAAGCCGATATTCTC 564455
QY 121 LeuAenAlaGlyLeuGluValThrGluLeuGluValThrAsnGluGlyProlysSerLeu 140
Db 564454 GCTAATGTCGGCAGAAAACCGTAATAATCGAAGTAACAATTCAGGCGACGCCCAATT 564395
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
Db 564394 CAAATTGGCTCGCATTAACCATTTTGTGAACCAATATATGCTTAATTTGACCGCACT 564335
QY 161 LysAlaTyGlyLysArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
Db 564334 TTGGCAGCGGAGATGCGCTTAATGTTCCATCGGCAATCGGTGCGTTTGACCCGCT 564275
QY 181 GlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200
Db 564274 GAAGTGAATCAGTGAAATTAAGTTGCTTTTGGTGTAACCAATCATTTATGTTCCAT 564215
QY 201 GlyLeuValAsn 204
Db 564214 AATCAATTGAT 564203

RESULT 15
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6538289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
; Mark D. Adams
; Owen White
; Hamilton O. Smith
; J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
; the Haemophilus influenzae Rd Genome, Fragments
; thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville,
; STATE: MD
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3 1/2 inch diskette
; OPERATING SYSTEM: MS DOS v6.22
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/643,990A
; FILING DATE: 23-Aug-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,429
; FILING DATE: 1995-06-07
; APPLICATION NUMBER: 08/426,787
; FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB16P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-610-5790
; TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1830121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

```

```

Alignment Scores:
Pred. No.: 3,79e-47 Length: 1830121
Score: 467.00 Matches: 96
Percent Similarity: 61.61% Conservative: 42
Best Local Similarity: 42.86% Mismatches: 62
Query Match: 39.54% Indels: 24
DB: 4 Gaps: 3

US-09-904-994B-2 (1-226) x US-09-643-990A-1 (1-1830121)
QY 3 LeuThrProLysGluGlnGluLysPheLeuLeuTyTYrAlaGlyGluValAlaArgLys 22
Db 564868 TTAACCTCCAGAGAACAGAAAACGATGCTTTTCCTCGCGGCGGAATCTCGCGGCAAA 564809
QY 23 ArgLysAlaGluGlyLeuLysLeuAsnGlnProGluAlaIleAlaTyTlIleSerAlaHis 42
Db 564808 CCGCAAGCAGCGCGGCGTTTAAATTAACATATCCAGAAACTATTTGCTTATTTGCTAGTCAT 564749
QY 43 IleMetAspGluAlaArgArgGlyLysLysThrValAlaGlnLeuMetGluGlyMet 62
Db 564748 TTACAAAGGCGCAGCAAGAAAGAA---ATGAGTGTAGCGGAAGTCAATGCAATATGCGCA 564692
QY 63 HisPheLeuLysLysAspGluValMetProGlyValGlyAsnMetValProAspLeuGly 82
Db 564691 ACACTTTTAACCGTTGATGATGATCATGGAAGGTGTGGCGGAATGCTTCATGAAATCCAG 564632
QY 83 ValGluAlaThrPheProAspGlyThrLysLeuValThrValAsnThrProIleGlu--- 101
Db 564631 ATTGAAGCTACTTTCCCGCATGGCAGAACTTGTTACCGTGCATATATCAATCAGATTA 564572
QY 101 ----- 101
Db 564571 CCGAAGGTGGGCTTTAGCCCAAAATTAATATCATATGTTGGGCTTAAGCCCACTT 564512
QY 102 ---ProAspGluHisPheLysAlaGlyGluValLysPheGlyCysAspLysAspIleGlu 120
Db 564511 ACAAGAACAAAGATATGATCCAGCGCAATACCAATTA---GCCAAGCGATATTTCTC 564455
QY 121 LeuAenAlaGlyLeuGluValThrGluLeuGluValThrAsnGluGlyProlysSerLeu 140
Db 564454 GCTAATGTCGGCAGAAAACCGTAATAATCGAAGTAACAATTCAGGCGACGCCCAATT 564395
QY 141 HisValGlySerHisPheHisPhePheGluAlaAsnLysAlaLeuLysPheAspArgGlu 160
Db 564394 CAAATTGGCTCGCATTAACCATTTTGTGAACCAATATGCGGTGCGTTTGAAACCCGCT 564275
QY 161 LysAlaTyGlyLysArgLeuAspIleProSerGlyAsnThrLeuArgIleGlyAlaGly 180
Db 564334 TTGGCAGCGGAGATGCGCTTAATGTTCCATCGGCAATGCGGTGCGTTTGAAACCCGCT 564275
QY 181 GlnThrArgLysValGlnLeuIleProLeuGlyGlySerLysValIleGlyMetAsn 200
Db 564274 GAAGTGAATCAGTGAAATTAAGTTGCTTTTGGTGTAACCAATCATTTATGTTCCAT 564215
QY 201 GlyLeuValAsn 204
Db 564214 AATCAATTGAT 564203

```

Search completed: February 16, 2005, 00:59:42  
Job time : 905.667 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - protein search, using frame\_nzp model

Run on: February 15, 2005, 20:03:24; Search time 31.5 Seconds

(without alignments)  
13664.333 Million cell updates/sec

Title: US-09-904-994B-1

Perfect score: 5147  
Sequence: 1 rgragatcttcacacactc.....aaaaagtagaagccacagg 2883

## Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 648760

Minimum DB seq length: 0

Maximum DB seq length: 100

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Command line parameters:

-MODEL=frame\_nzp.model -DEV=xlh  
-Q=/cg2\_1/USPRO.spool/US09904994/runat\_11022005\_134936\_22927/app\_query.fasta\_1.3079  
-DB=Issued\_Patents\_AA -OPMT=factan -SUFFIX=rai -MINMATCH=0.1 -LOOFC=0  
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blosum62 -TRANS=numan40.cdt  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=pco -NOM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=100  
-USRR=US09904994.QCGN\_1.1-26@runat\_11022005\_134936\_22927 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

Issued\_Patents\_AA.\*  
1: /cg2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cg2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cg2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cg2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cg2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cg2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	281.5	5.5	100	1	US-08-967-513-3
2	281.5	5.5	100	2	US-08-687-645B-3
3	269.5	5.2	99	1	US-07-732-242C-1
4	261.5	5.1	100	2	US-08-467-822-23
5	261.5	5.1	100	3	US-08-432-697-23
6	261.5	5.1	100	3	US-08-466-248-23
7	250.5	4.9	100	4	US-09-602-777A-18
8	87.5	1.7	92	4	US-09-252-991A-17446
9	78.5	1.5	96	4	US-09-328-352-4313
10	77	1.5	15	4	US-09-928-081-1
11	77	1.5	15	4	US-09-338-920B-1
12	71.5	1.4	93	4	US-09-489-039A-7356

C 13	70	1.4	89	4	US-09-489-039A-12832	Sequence 12832, A
C 14	69	1.3	99	4	US-09-328-352-4788	Sequence 4788, Ap
C 15	68.5	1.3	40	4	US-09-602-777A-12	Sequence 12, Appl
C 16	68	1.3	89	4	US-09-252-991A-17666	Sequence 17666, A
C 17	67.5	1.3	65	4	US-09-513-999C-7491	Sequence 7491, Ap
C 18	67	1.3	61	4	US-09-270-767-40422	Sequence 40422, A
C 19	67	1.3	61	4	US-09-270-767-55638	Sequence 55638, A
C 20	67	1.3	84	4	US-09-270-767-57094	Sequence 57094, A
C 21	67	1.3	96	4	US-09-513-999C-6065	Sequence 6065, Ap
C 22	67	1.3	99	2	US-08-537-400-19	Sequence 19, Appl
C 23	66.5	1.3	89	4	US-09-107-532A-3751	Sequence 3751, Ap
C 24	66.5	1.3	97	4	US-09-252-991A-27707	Sequence 27707, A
C 25	66	1.3	96	4	US-09-270-767-16576	Sequence 36576, A
C 26	66	1.3	96	4	US-09-270-767-51793	Sequence 51793, A
C 27	65.5	1.3	62	4	US-09-621-976-4501	Sequence 4501, Ap
C 28	65	1.3	75	4	US-09-471-780C-126	Sequence 5840, Ap
C 29	65	1.3	84	4	US-09-471-276-1347	Sequence 1347, Ap
C 30	64.5	1.3	81	5	PCT-US95-06266-126	Sequence 126, App
C 31	64	1.2	15	3	US-09-091-001-2	Sequence 2, Appl
C 32	64	1.2	95	4	US-09-252-991A-25355	Sequence 25355, A
C 33	63.5	1.2	67	1	US-08-471-780C-126	Sequence 126, App
C 34	63.5	1.2	67	1	US-08-467-282B-126	Sequence 126, App
C 35	63.5	1.2	67	2	US-08-471-282A-126	Sequence 126, App
C 36	63.5	1.2	67	2	US-08-466-710C-126	Sequence 126, App
C 37	63.5	1.2	67	3	US-08-468-739C-126	Sequence 126, App
C 38	63.5	1.2	67	4	US-09-293-769A-126	Sequence 126, App
C 39	63.5	1.2	67	4	US-08-537-871A-9	Sequence 9, Appl
C 40	63.5	1.2	74	3	US-08-679-493A-33	Sequence 33, Appl
C 41	63	1.2	63	4	US-09-134-000C-6110	Sequence 6110, Ap
C 42	63	1.2	74	4	US-09-270-767-34440	Sequence 34440, A
C 43	63	1.2	74	4	US-09-270-767-49657	Sequence 49657, A
C 44	63	1.2	78	4	US-09-949-016-10753	Sequence 10753, A
C 45	63	1.2	86	4	US-09-252-991A-29383	Sequence 29383, A

## ALIGNMENTS

RESULT 1  
US-08-967-513-3  
; Sequence 3, Application US/08967513  
; Patent No. 5783436  
; GENERAL INFORMATION:  
; APPLICANT: Robert P. Hausinger  
; TITLE OF INVENTION: Mutant Urease and Method  
; TITLE OF INVENTION: of Use For Determination  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ian C. McLeod  
; STREET: 2190 Commons Parkway  
; CITY: Okemos  
; STATE: Michigan  
; COUNTRY: USA  
; ZIP: 48864  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
; MEDIUM TYPE: storage  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/967, 513  
; FILING DATE: 11-NOV-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/687, 645  
; FILING DATE: July 26, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ian C. McLeod  
; REGISTRATION NUMBER: 20, 931  
; REFERENCE/DOCKET NUMBER: MSU 4.1-309  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (517) 347-4100

TELEFAX: (517) 347-4103  
TELEX: No. 5783436e  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100  
TYPE: amino acids  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: Protein  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: N/A  
STRAIN: N/A  
INDIVIDUAL ISOLATE: N/A  
CELL TYPE: N/A  
FEATURE:  
NAME/KEY: subunit Urea  
LOCATION:  
IDENTIFICATION METHOD: Sequencing  
OTHER INFORMATION: encoded subunit of  
OTHER INFORMATION: mutant urease  
US-08-967-513-3

Alignment Scores:  
Pred. No.: 8,43e-19 Length: 100  
Score: 281.50 Matches: 59  
Percent Similarity: 77.00% Conservative: 18  
Best Local Similarity: 59.00% Mismatches: 22  
Query Match: 5.47% Indels: 1  
DB: 1 Gaps: 1

US-09-904-994b-1 (1-2883) x US-08-967-513-3 (1-100)

QY 206 GTGAACCTCACACCCCAAGAGCAAGAAAGTTCTGTATATATATGCGGCGAAGTGCT 265  
Db 1 MetGluLeuThrProArgGluLysAspLysLeuLeuLeuPheThrAlaLeuValAla 20  
QY 266 AGAAGCCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325  
Db 21 GluArgArgLeuAlaArgGlyLeuLysLeuAsnTyrProGluSerValAlaLeuIleSer 40  
QY 326 GCCCATATTATGAGCAGACGCGCGCTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385  
Db 41 AlaPheIleMetGluGlyAlaArgAspGly--LysSerValAlaSerLeuMetGluGlu 59  
QY 386 TGCATGCACCTTTTGAAGAAAGATGAATATGCCCGGGGTGGTAAATATGTTCCCGAT 445  
Db 60 GlyArgHisValLeuThrArgGluGluGlnValMetGluGlyValProGluMetIleProAsp 79  
QY 446 CTAGGTGTGAAGCCACCTTTCTGTATGTATGTAACAACTTGTAATGATGGCCCATC 505  
Db 80 IleGlnValGluAlaThrPheProAspGlySerLysLeuValThrValHisAsnProIle 99

RESULT 2  
US-08-687-645B-3  
Sequence 3, Application US/08687645B  
Patent No. 5846752  
GENERAL INFORMATION:  
APPLICANT: Robert P. Hausinger  
TITLE OF INVENTION: Mutant Urease and Method  
TITLE OF INVENTION: of Use For Determination  
TITLE OF INVENTION: of Urea  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Ian C. McLeod  
STREET: 2190 Commons Parkway  
CITY: Okemos  
STATE: Michigan  
COUNTRY: USA  
ZIP: 48864  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Wordperfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/687,645B  
FILING DATE: July 26, 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ian C. McLeod  
REGISTRATION NUMBER: 20,931  
REFERENCE/DOCKET NUMBER: MSU 4.1-309  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (517) 347-4100  
TELEFAX: (517) 347-4103  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100  
TYPE: amino acids  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Protein  
DESCRIPTION: Protein  
HYPOTHETICAL: No  
ANTI-SENSE: No  
ORIGINAL SOURCE:  
ORGANISM: N/A  
STRAIN: N/A  
INDIVIDUAL ISOLATE: N/A  
CELL TYPE: N/A  
FEATURE:  
NAME/KEY: subunit Urea  
LOCATION:  
IDENTIFICATION METHOD: Sequencing  
OTHER INFORMATION: encoded subunit of  
OTHER INFORMATION: mutant urease  
US-08-687-645B-3

Alignment Scores:  
Pred. No.: 8,43e-19 Length: 100  
Score: 281.50 Matches: 59  
Percent Similarity: 77.00% Conservative: 18  
Best Local Similarity: 59.00% Mismatches: 22  
Query Match: 5.47% Indels: 1  
DB: 2 Gaps: 1

US-09-904-994b-1 (1-2883) x US-08-687-645B-3 (1-100)

QY 206 GTGAACCTCACACCCCAAGAGCAAGAAAGTTCTGTATATATATGCGGCGAAGTGCT 265  
Db 1 MetGluLeuThrProArgGluLysAspLysLeuLeuLeuPheThrAlaLeuValAla 20  
QY 266 AGAAGCCGCAAGCAGAGGGCTTAAAGCTCAACCAACCCGAAGCCATTGCTTACATTAGT 325  
Db 21 GluArgArgLeuAlaArgGlyLeuLysLeuAsnTyrProGluSerValAlaLeuIleSer 40  
QY 326 GCCCATATTATGAGCAGACGCGCGCTGGAAAAAACCCTTGCCAGCTTATGGAAGAG 385  
Db 41 AlaPheIleMetGluGlyAlaArgAspGly--LysSerValAlaSerLeuMetGluGlu 59  
QY 386 TGCATGCACCTTTTGAAGAAAGATGAATATGCCCGGGGTGGTAAATATGTTCCCGAT 445  
Db 60 GlyArgHisValLeuThrArgGluGluGlnValMetGluGlyValProGluMetIleProAsp 79  
QY 446 CTAGGTGTGAAGCCACCTTTCTGTATGTATGTAACAACTTGTAATGATGGCCCATC 505  
Db 80 IleGlnValGluAlaThrPheProAspGlySerLysLeuValThrValHisAsnProIle 99

```

RESULT 3
US-07-732-242C-1
; Sequence 1, Application US/07732242C
; Patent No. 528399
; GENERAL INFORMATION:
; APPLICANT: Iozumi, Takeshi; Masaki, Haruhiko;
; APPLICANT: Hidaka, Makoto; Nakamura, Akira;
; APPLICANT: Maeda, Michihisa; Yonets, Yasuo
; TITLE OF INVENTION: Gene of Urease
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Frieheuf, Holtz, Goodman & Woodward, P.C.
; STREET: 600 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10016-2088
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.72mb
; COMPUTER: IBM PC compatible (NEC PC-9801 RX)
; OPERATING SYSTEM: MS DOS
; SOFTWARE: ASCII form
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/732,242C
; FILING DATE: 19910718
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JPN 2-210178
; FILING DATE: 10-AUG-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Herbert
; REGISTRATION NUMBER: 17081
; REFERENCE/DOCKET INFORMATION:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 972-1400
; TELEFAX: (212) 370-1622
; TELEX: 236268
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-732-242C-1

Alignment Scores:
Pred. No.: 1,25e-17 length: 99
Score: 269.50 Matches: 55
Percent Similarity: 73.74% Conservative: 18
Best Local Similarity: 55.56% Mismatches: 25
Query Match: 5.24% Indels: 1
DB: 1 Gaps: 1

US-09-904-994B-1 (1-2883) x US-07-732-242C-1 (1-99)

QY 209 AACTCACACCCAGCAAGAAGATTCTTGTTATATTATGCGGGGAAGTGTAGTA 268
| | | | | : : : | | | | | : : : | | | | |
Db 1 LysleuthrSerArgIleuMetGluLysleuMetIleValAlaIleAspLeuAlaArg 20
QY 269 AAGCGCAAGACAGAGGGCTTAAAGTCACCAACCCGAGACCATTGCTTACATTAGTGC 328
: : : | | | | | : : : | | | | | : : : | | | | |
Db 21 ArgArgLysGluArgGlyLeuLysleuAsnTyrProGluValAlaIleMetIleThrTyr 40
QY 329 CATATTATGACGAGAGCGGCGGTGGAAAAAACCGTTGCCAGCTTATGAGAAAGTGC 388
: : : : : | | | | | | | | | | | | | | | | | | | | |
Db 41 GluValIleGlnGlyValIleArgSerpGly--LysThrValIleGlnLeuMetGlnTyrGly 59
QY 389 ATGCACTTTTGAATAAAGATGAAGTAATGCCCGGGGTGGGTATATGTTCCCGATCTA 448
| | | | | : : : : : | | | | | | | | | | | | | | | |
Db 60 AlaThrIleLeuThrLysGluAspValIleMetGlnGlyValAlaGlnMetIleProAspIle 79
449 GGTTAGAAAGCACTTCTTCGTATGTTAGCAAACTTGTAACCTGTGAATTGGCCCATC 505

```

```

9b      80 GlnIleGlnIleValThrProAspGlyThrIleValThrValHisAspProIle 98
      RESULT 4
      US-08-467-822-23
      Sequence 23, Application US/08467822
      Patent No. 5843460
      GENERAL INFORMATION:
      APPLICANT: Labigne, Agnee
      APPLICANT: Sauterbaum, Sebastien
      APPLICANT: Ferrero, Richard L.
      APPLICANT: Thiberge, Jean-Michel
      TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
      HELICOBLACTER INFECTION POLYPEPTIDES FOR USE IN THE
      TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
      NUMBER OF SEQUENCES: 44
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Fimnegan, Henderson, Farabow, Garrett &
      STREET: 1300 I Street, N.W.
      CITY: Washington
      STATE: D.C.
      COUNTRY: USA
      ZIP: 20005-3115
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/467,822
      FILING DATE: 06-JUN-1995
      CLASSIFICATION: 435
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/447,177
      FILING DATE: 19-MAY-1995
      CLASSIFICATION: 435
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 08/432,697
      FILING DATE: 02-MAY-1995
      CLASSIFICATION: 435
      ATTORNEY/AGENT INFORMATION:
      NAME: Meyers, Kenneth J.
      REGISTRATION NUMBER: 25,146
      REFERENCE/DOCKET NUMBER: 03495_0137-02000
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (202) 408-4000
      TELEFAX: (202) 408-4400
      INFORMATION FOR SEQ ID NO: 23:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 100 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      US-08-467-822-23
      Alignment Scores:
      Score: 7.61e-17 Length: 100
      Pident: 261.50 Matches: 56
      Percent Similarity: 73.00% Conservative: 17
      Best Local Similarity: 56.00% Mismatches: 26
      Query Match: 5.08% Indels: 1
      Gaps: 1
      US-09-904-994B-1 (1-2883) x US-08-467-822-23 (1-100)
      206 GTGAAGTACACCCCAAGAGCAAGAAAGTTCTTTATATATATGCGGCGAGTGCT 265
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
      Db 1 MecGlueuethrProrAgtIuIySaapIySleuIeuIeuPhetThrAlaGlyIeuValAla 20
      266 AGAAAGCGCAAGACAGAGGGCTTAAAGTCAACCAACCCGAAGCATTGCTTACATTAGT 325

```

Db 21 GUARGARGLeuAlaIysGlyLeuIbSLeuAenTyProGluArgValAlaLeuIleSer 40  
QY 326 GCCCATATATGAGCAAGAGCGCCGCTGGAAAAAACCCTTGCCACGCTTATGGAAGAG 385  
Db 41 CysAlaIleMetGluGlyAlaArgGluGly---LysThrValAlaGlnLeuMetSerGln 59  
QY 386 TGCATGCACTTTTGGAAAAAGATGAAGTAATGCCGGGGGTGGTAATATGTTCCCGAT 445  
Db 60 GYAAGThValLeuThrAlaGluGlnValMetGluGlyValProGluMetIleLysAsp 79  
QY 446 CTAGGTGTAAGAACCACTTCTCGATGTACGAAACTGTGAATGGAATGGCCCATC 505  
Db 80 ValGlnValGluCysThrPheProAspGlyThrLysLeuValSerIleHisSerProIle 99

RESULT 5  
US-08-432-697-23  
Sequence 23, Application US/08432697  
Patent No. 6248330

## GENERAL INFORMATION:

APPLICANT: Labigne, Agnes  
APPLICANT: Sauerbaum, Sebastien  
APPLICANT: Ferrero, Richard L.  
APPLICANT: Thiberge, Jean-Michel  
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
TITLE OF INVENTION: HELICOBLACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/432,697  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495.0137-00000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-432-697-23

## Alignment Scores:

Pred. No.: 7.61e-17 Length: 100  
Score: 261.50 Matches: 56  
Percent Similarity: 73.00% Conservative: 17  
Best Local Similarity: 56.00% Mismatches: 26  
Query Match: 5.08% Indels: 1  
DB: 3 Gaps: 1

US-09-904-994b-1 (1-2883) x US-08-432-697-23 (1-100)

QY 206 GTGAACCTGACACCAAGAGCAAGAAAGTTCTGTATATTATATGCGGGCGAAGTGCGT 265

Db 1 MetGluLeuThrProArgGluLysAspLysLeuLeuLeuPheThrAlaGlyLeuValAla 20  
QY 266 AAAAAAGCGAAGCAGAGGGCTTAAAGCTCAACCAAGCAAGCCGAACTTCTTACATTACT 325  
Db 21 GUARGARGLeuAlaIysGlyLeuIbSLeuAenTyProGluArgValAlaLeuIleSer 40  
QY 326 GCCCATATATGAGCAAGAGCGCCGCTGGAAAAAACCCTTGCCACGCTTATGGAAGAG 385  
Db 41 CysAlaIleMetGluGlyAlaArgGluGly---LysThrValAlaGlnLeuMetSerGln 59  
QY 386 TGCATGCACTTTTGGAAAAAGATGAAGTAATGCCGGGGGTGGTAATATGTTCCCGAT 445  
Db 60 GYAAGThValLeuThrAlaGluGlnValMetGluGlyValProGluMetIleLysAsp 79  
QY 446 CTAGGTGTAAGAACCACTTCTCGATGTACGAAACTGTGAATGGAATGGCCCATC 505  
Db 80 ValGlnValGluCysThrPheProAspGlyThrLysLeuValSerIleHisSerProIle 99

RESULT 6  
US-08-466-248-23  
Sequence 23, Application US/08466248  
Patent No. 6258359

## GENERAL INFORMATION:

APPLICANT: Labigne, Agnes  
APPLICANT: Sauerbaum, Sebastien  
APPLICANT: Ferrero, Richard L.  
APPLICANT: Thiberge, Jean-Michel  
TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST  
TITLE OF INVENTION: HELICOBLACTER INFECTION, POLYPEPTIDES FOR USE IN THE  
TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID  
TITLE OF INVENTION: POLYPEPTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESS: Dunner  
STREET: 1300 I Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-3315

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,248  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/447,177  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/432,697  
FILING DATE: 02-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Meyers, Kenneth J.  
REGISTRATION NUMBER: 25,146  
REFERENCE/DOCKET NUMBER: 03495.0137-02000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 408-4000  
TELEFAX: (202) 408-4400  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-466-248-23

Alignment Scores:  
Pred. No.: 7.61e-17 Length: 100  
Score: 261.50 Matches: 56  
Percent Similarity: 73.00% Conservative: 17  
Best Local Similarity: 56.00% Mismatches: 26  
Query Match: 5.08% Indels: 1  
DB: 3 Gaps: 1

US-09-904-994B-1 (1-2883) x US-08-466-248-23 (1-100)

QY 206 GTGAACTCACACCCAGCAAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCGT 265  
Db 1 MetcIuLeuthrProAArgGluYAspIysLeuLeuLeuPheTrAlaGlyLeuValAla 20

QY 266 AGAAAGCCGCAAGAGAGGCGTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGT 325  
Db 21 GluArgArgLeuAlaIalysGlyLeuLeuLeuLeuTyrProGluArgValAlaLeuIleSer 40

QY 326 GCCCATTTTATGACGAAAGCGCGCGTGGAAAAAACCCTTGCCCGCATTATGGAAGAG 385  
Db 41 CysAlaIleMetGluGlyAlaArgGluGly--LysThrValAlaGlnLeuMetSerGlu 59

QY 386 TGCATGCACTTTTGAAGAAAGATGATGCGCGGCGGTGAATATGATGTTCCCGAT 445  
Db 60 GlYArgThrValIleuThrAlaGlnValMetGluGlyAlaProGluMetIleLysAsp 79

QY 446 CTAGGTGTAGAAAGCCACTTCTGTATGTATGCAAGAACTTGAATGTGAATGGCCCATC 505  
Db 80 ValGlnValGluCysThrPheProAspIlyThrIysLeuValSerIleHisSerProIle 99

RESULT 7  
US-09-602-777A-18  
; Sequence 18, Application US/09602777A  
; Patent No. 6831165  
; GENERAL INFORMATION:  
; APPLICANT: Pompeius, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schröder, Hartwig  
; APPLICANT: Zelder, Oskar  
; APPLICANT: Haberhauser, Gregor  
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS  
; FILE REFERENCE: BGI-1286P  
; CURRENT APPLICATION NUMBER: US/09/602, 777A  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/141031  
; PRIOR FILING DATE: 1999-06-25  
; PRIOR APPLICATION NUMBER: DE 19931636.8  
; PRIOR FILING DATE: 1999-07-08  
; PRIOR APPLICATION NUMBER: DE 19932125.6  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932126.4  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932127.2  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932128.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932129.9  
; PRIOR FILING DATE: 1999-07-19  
; PRIOR APPLICATION NUMBER: DE 19932226.0  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: DE 19932920.6  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932922.2  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932924.9  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932928.1  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932930.3  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932933.8

PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932935.4  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19932973.7  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933002.6  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933003.4  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933005.0  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19933006.9  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: DE 19941378.9  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941379.7  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941390.8  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19941391.6  
; PRIOR FILING DATE: 1999-08-31  
; PRIOR APPLICATION NUMBER: DE 19942088.2  
; PRIOR FILING DATE: 1999-09-03  
; NUMBER OF SEQ ID NOS: 442  
; SEQ ID NO 18  
; LENGTH: 100  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-602-777A-18

Alignment Scores:  
Pred. No.: 9.05e-16 Length: 100  
Score: 250.50 Matches: 51  
Percent Similarity: 72.45% Conservative: 20  
Best Local Similarity: 52.04% Mismatches: 26  
Query Match: 4.87% Indels: 1  
DB: 4 Gaps: 1

US-09-904-994B-1 (1-2883) x US-09-602-777A-18 (1-100)

QY 212 CTCACACCAAGCAAGCAAGAAAGTTCTGTATATATGCGGCGAAGTGCTAGAAG 271  
Db 3 IleThrProArgGluGlnGluYlYleuMetIleValValAlaIAspLeuAlaArg 22

QY 272 CGCAAGCAAGAGGCTTAAAGCTCAACCAACCCGAGCCATTGCTTACATTAGTGGCCAT 331  
Db 23 ArgYAspArgGlyLeuYlYleuAsnHleProGluAlaValAlaLeuIleThrGlu 42

QY 332 ATTATGACGAAAGCGCCGCTGAAAAAACCCTGCCAGCTTATGGAAGTGCATG 391  
Db 43 LeuIleGluGlyAlaArgAspGly--ArgThrValAlaAspLeuMetSerTrpGlySer 61

QY 392 CACTTTTGAAGAAAGATGAAGTATGCCCGGCGGTATATGTTCCCGATCTAGT 451  
Db 62 ThrIleLeuThrArgAspArgValLeuGluGlyIleProGluMetIleProAspIleGln 81

QY 452 GTAGAGCCACCTTCTGTATGTATGCAAGAACTTGAATGTGAATGGCCCATC 505  
Db 82 ValGlnAlaThrPheAspArgGlyThrIlyLeuValThrValHisAsnProIle 99

RESULT 8  
US-09-252-991A-17446  
; Sequence 17446, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252, 991A  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074, 788  
; PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: 60/094,190  
 PRIOR FILING DATE: 1998-07-27  
 NUMBER OF SEQ ID NOS: 33142  
 SEQ ID NO 17446  
 LENGTH: 92  
 TYPE: PRY  
 ORGANISM: *Pseudomonas aeruginosa*  
 US-09-255-991A-17446

Alignment Scores:	
Pred. No.:	7.43
Score:	87.50
Percent Similarity:	37.76%
Best Local Similarity:	25.17%
Query Match:	1.70%
DB:	4
	7
	Gaps:
	92
	Length:
	36
	Matches:
	18
	Conservative:
	24
	Mismatch:
	65
	Indels:
	7

US-09-904-994B-1 (1-2883) X US-09-252-991A-17446 (1-92)

OY	2101	AGCGCTCATCTCCAAATACACTATCAACC	CGGCTTTGA	CCACGAGCGGTAGCGAGTATA	2160
Dd	6	SerialNrSerProSerLeuSerAlaProbola	-----	AlaSerSer	20
OY	2161	TGGCGTCGTGGAAGAGGCCAAGAATCG	CCGCACTTG	TGNTGNGANTCCTGCTTTTTYG	2220
Dd	21	SerSerSerTrpLysArgSerAla	-----	-----	28
OY	2221	GCGTAAAAACCAAATCGTGATC	AAAGGCGGTATG	TGTCCTTCTGTAAAATGGCGCAT	2280
Dd	29	-----	-----	AlaAlaAlaIaArgTrpSerAla	35
OY	2281	CTAACGCGTCTGTGC	CCCACTCCCA	CCGGTTATTATCACGGAATGTTTGCGCAT	2340
Dd	36	-----	-----	ThrArgCyMet	39
OY	2341	GCAAGGCGAAATTGCACACGACATCA	CTTTGTTTCCA	AAGTGGCTATGAAATGGCG	2400
Dd	40	-----	-----	SerAlaIleSerCySerSerPrAlaIleSerAla	51
OY	2401	TGAAG-----	AAAGCTGGGCTTA	GAGCGCCAAAGTTCTACCGGTCAAAAAC	2451
Dd	52	AlaLysAsnArgProArgSerTrpArg	-----	ArgSerCyTherSer	65
OY	2452	GTAACATCACCAAGAAAGACTTCAAG	TTCAAAGACA	CAAAACGGCAAAAATCACCGTGCAT	2511
Dd	66	AlaSerSerProGlyLysThrSerAlaentrp	-----	ArgArgAlaSerProLysIle	83
OY	2512	CGAAAAACCT	2520	-----	-----
Dd	84	ArgValPro	86	-----	-----

```

RESULT 9
US-09-328-352-4313
; Sequence 4313; Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4313
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4313

```

Alignment Scores:	
Pred. No.:	57.8
Score:	78.50
Percent Similarity:	41.35%
	Length: 96
	Matches: 22
	Conservative: 21

Best Local Similarity:	21.15%	Mismatches:	32
Query Match:	1.53%	Indels:	29
DB:	4	Gaps:	3

US-09-904-994B-1 (1-2883) X US-09-328-352-4313 (1-96)

```

Qy 1381 GGCACCTGTGGCGCGTCCCAACATGTTGTAAACGCCATTGGCTGAGCGGTAGGGAATGT 1322
    ||| ||||| ||||| |||
Db 9 GlyAspAlaProProPro----- 14

Qy 1321 TGTGAGAAAGAAAGTGGGTGTGTGAATCGATTCCCGGTAATATCATCTTCC 1262
    :: :: :: ||| ||| :: ::
Db 15 -----11MetProProValThrProProIleIleSerValThrThr 28

Qy 1261 CTGCTAGTGCCTTCTGTGGCCACACCCACGACCATATGAGGCGCTTACGCACTTTCATG 1202
    ||| ||| ||| ||| ||| ||| ||| |||
Db 29 ProArgIleuSerProProThrProPro-----ValThrProProIleIle 44

Qy 1201 TCCTTGTGTTCTGCCTTGCCATGCCATGATGATTTTGGCGTTTAAATCCCATGTCGACT 1142
    ||| :: :: ||| ||| ||| ||| ||| |||
Db 45 SerValIleThrSerSerIleuProSerThrProProValThrProProIleThrSer 64

Qy 1141 TTGTAAATCCCGGTGTGATCGATN-----ATCATCGCGTTAGTGATGACT 1097
    :: :: ||| :: :: :: ||| ||| ||| |||
Db 65 ValIleuProAlaThrProValThrIleProValAlaValAlaProPheValIleSer 84

Qy 1096 AAATCTAGGGTG 1085
    ||| ||| ||| |||
Db 85 ValSerArgIle 88

```

RESULT 10

US-08-928-081-1  
; Sequence 1, Application US/08928081

```

1 GENERAL INFORMATION:
2 APPLICANT: Soman, Gopalan
3 APPLICANT: Thomas, Jr., William D.
4 APPLICANT: Monath, Thomas P.
5 TITLE OF INVENTION: Stabilization of
6 NUMBER OF SEQUENCES: Helicobacter Urease
7 NUMBER OF SEQUENCES: 4
8 CORRESPONDENCE ADDRESS:
9 ADDRESS: Clark & Elbing LLP
10 STREET: 176 Federal Street
11 CITY: Boston
12 STATE: MA
13 COUNTRY: USA
14 ZIP: 02110
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Diskette
17 OPERATING SYSTEM: IBM Compatible
18 SOFTWARE: FASTSEQ for Windows Version 2.0
19 CURRENT APPLICATION DATA:
20 APPLICATION NUMBER: US/08/928,081

```

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 06133/0230001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

```

/ INFORMATION FOR SEQ ID NO: 1
/
/ SEQUENCE CHARACTERISTICS:
/
/ LENGTH: 15 amino acids
/
/ TYPE: amino acid
/
/ STRANDEDNESS: single
/
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: peptide
/
US-08-928-081-1

```

## Alignment Scores:

Pred. No.: 27 Length: 15  
Score: 77.00 Matches: 13  
Percent Similarity: 93.33% Conservative: 1  
Best Local Similarity: 86.67% Mismatches: 1  
Query Match: 1.50% Indels: 0  
DB: 2 Gaps: 0

US-09-904-994B-1 (1-2883) x US-08-928-081-1 (1-15)

QY 1527 GAAGCGGCGCGATTGTTTAAATTCATGAGAGCTGGGCGACA 1571

Db 1 GluAaGlyAlaIleGlyPheAlaIleHisGluAspTrpGlyThr 15

## RESULT 11

US-09-338-920B-1  
Sequence 1, Application US/09338920B  
Patent No. 6703851  
GENERAL INFORMATION:

APPLICANT: Soman, Gopalan  
APPLICANT: Thomas, William D., Jr.

APPLICANT: Monath, Thomas P.

TITLE OF INVENTION: Stabilization of Helicobacter Urease

FILE REFERENCE: 06132/023002

CURRENT APPLICATION NUMBER: US/09/338,920B

PRIOR FILING DATE: 1997-09-12

NUMBER OF SEQ ID NOS: 14

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1

LENGTH: 15

TYPE: PRT

ORGANISM: Helicobacter pylori

## Alignment Scores:

Pred. No.: 27 Length: 15  
Score: 77.00 Matches: 13  
Percent Similarity: 93.33% Conservative: 1  
Best Local Similarity: 86.67% Mismatches: 1  
Query Match: 1.50% Indels: 0  
DB: 4 Gaps: 0

US-09-904-994B-1 (1-2883) x US-09-338-920B-1 (1-15)

QY 1527 GAAGCGGCGCGATTGTTTAAATTCATGAGAGCTGGGCGACA 1571

Db 1 GluAaGlyAlaIleGlyPheAlaIleHisGluAspTrpGlyThr 15

## RESULT 12

US-09-489-039A-7356  
Sequence 7356, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 7356

LENGTH: 93

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

## Alignment Scores:

Pred. No.: 274 Length: 93

Score: 71.50 Matches: 32

Percent Similarity: 37.27% Conservative: 9

Best Local Similarity: 29.09% Mismatches: 46

Query Match: 1.39% Indels: 23

DB: 4 Gaps: 6

US-09-904-994B-1 (1-2883) x US-09-489-039A-7356 (1-93)

QY 2026 TGCCAAAGTTCGAGCAATCCTCGCTTCA-----CGCCCC-----ATTGCTTGC 1982

Db 2 CysArgSerThrSerAlaValAlaProCysSerThrIleArgProAlaTyrIleThr 21

QY 1981 GAATCCAGCTTTCATCGATCGATCACCATTATCATGAGACATCTTCAGCCGCGATA 1922

Db 22 ThrCysSerAlaIleCysAlaIleThrProArgSerTrpValIleSerIleIleAla 41

QY 1921 GAGCCGGGGCGGATACGCTTTTGAGAAATTTGAAATCTCGCGGATGCGTTGCTAGG 1862

Db 42 TrpLeuAlaArgAlaArg-----SerArgSerArgArgArgIle----- 54

QY 1861 TGGTGCATGTGATGACATGTCTAAGTCTTCGCAACCGTATTATGATAGGATA 1802

Db 55 -----CysAlaCysThrValThrSerSerAlaValAlaGlySer 67

QY 1801 GTGGGGGTGGGAGAGAGGAGGAGATATATGAGCTGCGCTGCATGATGATACATGAGT 1742

Db 68 SerAlaIleSerSerLeuGlyArgGlnSerAlaMetAlaIleIleThr----- 84

QY 1741 GAGTGTCTCCACCCGCTCCCTCAATGTGG 1712

Db 85 ArgCysSerMetPro---ProGluSerTrp 93

## RESULT 13

US-09-489-039A-12832  
Sequence 12832, Application US/09489039A  
Patent No. 6610836  
GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 12832

LENGTH: 89

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-904-994B-1 (1-2883) x US-09-489-039A-12832 (1-89)

QY 645 TGGAAATGCTACCCACATGCAAGATTAGGC---CCTTCATTAGTAACCTCAAGTTGC 589

Db 2 TrpIleThrIleAlaIleThrProArgArgAlaGlyThrThrSerPheIleThrSerPro 21

QY 588 GTTACTCTTTGGCTGATGAGCTGATGCTTTATCGAACAATTCATTCAGCTGGCCC 529

Db 22 IleSerThrSerProSerLeuThrSer-----SerSerPro 33

QY 528 GCTTTGAAGTCTCATCTGCTGATGAGGCGCAATTACAGTTCAAGTTTGTACATCA 469

Db 34 AlaIleMetArgSerSerValAlaPheProGln-----ProGlu 46

```
QY 468 GGA-----AAGTGCTTCTACACTATA-----TCGGA 439
Db 47 GlyProThrLysThrIleAsnSerIleSerArgLeuThrTrpCysSerVal 66
QY 438 ACCATATACCAACCCCGGCAATTACT 412
Db 67 ThrValLeuProAsnLeuLeuValThr 75

RESULT 14
US-09-328-352-4788
; Sequence 4788, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 4788
; LENGTH: 99
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-4788

Alignment Scores:
Pred. No.: 499 Length: 99
Score: 69.00 Matches: 33
Percent Similarity: 32.65% Conservative: 15
Best Local Similarity: 22.45% Mismatches: 37
Query Match: 1.34% Indels: 62
DB: 4 Gaps: 8

US-09-904-994b-1 (1-2883) x US-09-328-352-4788 (1-99)
QY 1573 GTTGGCCCCAGCTTCATGCAATTAAACCAATGCGCGCGCTTCTACTTGTCTACA 1514
Db 1 IleIleProValSer-----GlyLeuSerProArgSerPro----- 12
QY 1513 AGTTGTTTTTGTAGAGTATTGCTTTGCCCAAAAGCCCATTCATGAGATATCTG 1454
Db 13 -----PrometProProMetProProIleMetPro----- 22
QY 1453 TCTGCTCGCGCAACATGCGGTGCAAGTTCATTCGCCGAGATGATAGTCGATTC 1394
Db 23 -----ValIleGlyLeuLysIle 28
QY 1393 GTGCCATCTACAGACCTGCGCGCTT-----CCA 1364
Db 29 SerProSerPro---ProIleProProMetIleProValIleGlyLeuLysIleSerPro 47
QY 1363 AACATGTTGTAAACGCTATTGCTAGACGGTAAAGGAATTGTTGTGAGAAAGAGTGG 1304
Db 48 SerProProIleProProIleThrProValIleGly----- 59
QY 1303 GTGTGTGAATCCATTCCTCCAGGGGTAATATATACCTTCCCTCGTAGTGTCTGTG 1244
Db 60 -----ProLysIleSerProValProSerProProIleAsnGluLys 73
QY 1243 CCCAGCCACGACGACATGAGGGCTT-----ACGCACTTGCATGTCCTTG 1196
Db 74 GluLeuProAsnThrLeuIleGlyValIleGlyAlaThrThrSerThrSerPro 93
QY 1195 TTTCCTGCTTGGCAATGCCA 1175
Db 94 LeuPro---LeuProMetPro 99

RESULT 15
US-09-602-777A-12
; Sequence 12, Application US/09602777A
; Patent No. 6831165
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Pompejus, Markus
/ APPLICANT: Kroege, Burkhard
/ APPLICANT: Schroeder, Hartwig
/ APPLICANT: Zelder, Oskar
/ APPLICANT: Haberbauer, Gregor
/ TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
/ INVOLVED IN HOMEOSTASIS AND ADAPTATION
/ FILE REFERENCE: BGI-128CP
/ CURRENT APPLICATION NUMBER: US/09/602,777A
/ CURRENT FILING DATE: 2000-06-23
/ PRIOR APPLICATION NUMBER: US 60/141031
/ PRIOR FILING DATE: 1999-06-25
/ PRIOR APPLICATION NUMBER: DE 19931636.8
/ PRIOR FILING DATE: 1999-07-08
/ PRIOR APPLICATION NUMBER: DE 19932125.6
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932126.4
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932127.2
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932128.0
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932129.9
/ PRIOR FILING DATE: 1999-07-19
/ PRIOR APPLICATION NUMBER: DE 19932226.0
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: DE 19932920.6
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932922.2
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932924.9
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932928.1
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932930.3
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932933.8
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932935.4
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19932973.7
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19933002.6
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19933003.4
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19933005.0
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19933006.9
/ PRIOR FILING DATE: 1999-07-14
/ PRIOR APPLICATION NUMBER: DE 19941378.9
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19941379.7
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19941390.8
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19941391.6
/ PRIOR FILING DATE: 1999-08-31
/ PRIOR APPLICATION NUMBER: DE 19942088.2
/ PRIOR FILING DATE: 1999-09-03
/ NUMBER OF SEQ ID NOS: 442
/ SEQ ID NO 12
/ LENGTH: 40
/ TYPE: PRN
/ ORGANISM: Corynebacterium glutamicum
US-09-602-777A-12

Alignment Scores:
Pred. No.: 327 Length: 40
Score: 68.50 Matches: 14
Percent Similarity: 48.89% Conservative: 8
Best Local Similarity: 31.11% Mismatches: 16
```



Query Match: 1.33% Indels: 7  
DB: 4 Gaps: 1

US-09-904-994B-1 (1-2883) x US-09-602-777A-12 (1-40)

```
QY 509 CCAGATGAGCACTTCAAAGCGGCGAAGTGAATTGGTTGCGATAAAGACATGAGCTC 568
    ||| |||::: ||| |||
Db 3 ProglyglutryrileuenserSercluserleuthrgly----- 15
QY 569 AATGCAGGCAAGAGTAACCGAAGTGTACTAATGAGGCGCTAAATCCTTGAT 628
    ||| |||::: ||| |||
Db 16 AsnValGlyArgGluAlaLysThrIleGluIleAsnThrGlyAspArgProValGln 35
    ||| |||::: ||| |||
QY 629 GTGGTAGCCATTTC 643
    :::||| ||| |||
Db 36 IleGlySerHisPhe 40
    :::||| ||| |||
```

Search completed: February 15, 2005, 20:34:27  
Job time : 36.5 secs

**This Page Blank (uspto)**

**This Page Blank (uspto)**